Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis

IMSGC

Table of Contents

ImmunoChip Details

Basic content

ImmunoChip is an Illumina Infinium HD custom array designed to enable fine mapping of established autoimmune loci and deeply replicate autoimmune GWAS results.[1] In total 207728 variants were submitted for inclusion on the chip; of which 196524 passed manufacturing quality control and were ultimately included on the chip (192402 autosomal, 1595 X-linked, 1735 Y-linked, 791 pseudoautosomal and one mitochondrial). In its final form the chip included 189 non-MHC fine mapping regions, two on the X chromosome and the other 187 on the autosomes. Three of the autosomal regions included less than 10 usable markers; leaving 184 deeply interrogated non-MHC fine mapping regions on the chip.

The multiple sclerosis contribution to the ImmunoChip was based on our meta-analysis of the published GWAS,[2-5] their related follow-up efforts[6-9] and a preliminary analysis of the GWAS we performed in collaboration with WTCCC2.[10] Based on these data we nominated a total of 26 regions for fine mapping (including the MHC); 15 of these were uniquely nominated by multiple sclerosis, while the other 11 were also nominated by at least one other autoimmune disease group in the ImmunoChip consortium.

We also nominated 5887 SNPs for replication; 4794 identified from our meta-analysis of published GWAS[11] (including all 206 SNPs with a p value of $< 10^{-5}$ in that analysis) and a minimally overlapping set of 1130 identified from our preliminary analysis of the IMSGC and WTCCC2 GWAS[10] (including 643 independent signals and proxies for these where available). In total, 2212 of these were eventually included on the chip, of which 1917 passed QC, including 1012 of the 4794 originally nominated from the meta-analysis[11] and 935 of the 1130 originally nominated from the preliminary analysis of the IMSGC and WTCCC2 GWAS.[10] Some 481 of these are part of fine mapping intervals leaving 1436 outside these intervals; 776 from the meta-analysis[11] and 671 from the latest GWAS[10] (only 11 markers were nominated by both). Linkage disequilibrium (LD) pruning these sets to generate independent signals left 567 from the meta-analysis[11] and 317 from the GWAS[10] (with just one marker in common).

In addition a further 2109 SNPs of local interest (wild cards) were also nominated, 659 of these were eventually included on the chip and 394 passed QC. Some 110 of these were included in fine mapping regions. The remainder (284) included several regions of extensive LD, so after LD pruning there were just 52 independent signals. The multiple sclerosis related SNPs selected for inclusion on the ImmunoChip were picked at an early stage in the analysis of the main GWAS (and the meta-analysis of previous GWAS) with the result that not all of the 57 regions eventually reported in the GWAS have been included as fine mapped regions on ImmunoChip.

Eleven of the 935 SNPs nominated from the 2011 GWAS[10] are included amongst the 48 novel associated SNPs listed in Table 1 of the main text. A further 20 of the SNPs on Table 1 have proxies amongst the 935 (13 of which pass the $p < 1 \times 10^{-4}$ threshold in the discovery analysis). Overlap also exists with the results from the previous meta-analysis of GWAS,[11] with 5 of the SNPs from Table 1 having proxies amongst the 1012 SNPs nominated from that study. In total >50% (26/48) of the associations listed in Table 1 would have been captured by follow up study of just these MS specified sources.

Genomewide significant evidence for association was seen in multiple sclerosis in 25 of the 26 fine mapping regions nominated by the IMSGC and WTCCC2 (one region containing two independent signals of association). The only nominated region that failed to identify such evidence was one of the three fine mapping regions on the ImmunoChip where there were ultimately less than 10 markers passing QC. In total 59 of the 97 associations reaching genomewide significance in our ImmunoChip data lie in fine mapping regions (two regions each contain two signals of association and one region contains three such signals). Substantial enrichment for genuine association was also seen amongst the deep replication SNPs nominated by IMSGC and WTCCC2, as is demonstrated by the marked genomic inflation apparent in the QQ plots for these SNPs (those passing QC and lying outside the fine mapping regions) (see Supplementary Figure 1 and Supplementary Figure 2). The multiple sclerosis specified content on the ImmunoChip included proxies for all but 14 of the 48 novel associations reaching genome-wide significance.

Previously published primary associations

In our 2011 GWAS[10] we listed 57 primary associations - 23 that had been identified in earlier GWAS (and their related follow up efforts) and 34 that were novel (29 reaching genomewide significance and 5 that just missed this threshold; all of these 5 have subsequently been confirmed with genomewide significance[12]). We listed a further 3 different SNPs in our meta-analysis of previously published GWAS, [11] however two of these are now known to be tagging already reported signals (rs170934 and rs6718520) and the final SNP (rs2150702) has no good proxies included on the ImmunoChip.
Of the 57 primary SNPs from the 2011 GWAS[10] 51 were included on the ImmunoChip and good proxies for the other six were also included (see Supplementary Table 3). For every SNP the previously associated allele was again over represented in ImmunoChip cases and for all but 2 SNPs (rs2028597 and rs802734) the difference was at least nominally significant (one sided $p < 1.0 \times 10^{-1}$). Association with rs9657904 ($r^2 = 0.271$ and D` = 1 with rs2028597) in *CBLB* was originally reported in the Sardinian GWAS[13]. There has been modest replication in other populations such as the continental Italian population,[14] but never as strong. We conclude that the lack of signal in the ImmunoChip data set may be due to both an LD specific effect in the Sardinian population and a lack of adequate tagging. The SNP rs802734 has a D' of 1.0 with the nearby SNP rs9482848 (less than 2kb) which is modestly associated ($p = 1.0 \times 10^{-3}$) indicating that rs802734 is probably only modestly tagging the functionally relevant variant at this site.

Previously published secondary associations

In our original GWAS[10] conditional analysis based on the primary signals suggested the existence of 7 additional secondary associations, 4 of these replicated and were therefore reported - one (rs12048904) close to rs11581062 on Chr1, another (rs7090512) close to rs3118470 on Chr10 and two signals (rs4285028, rs4308217) close to rs9282641 on Chr3.

Chromosome 1p21 (in the region of *VCAM1*, *EXTL2* and *SLC30A7*)

In the original GWAS[10] the most associated (lead) SNP in this region was rs11581062. Conditioning on this lead SNP revealed association with a second SNP rs12048904; both of these signals replicated and gave genomewide significant association in the combined analysis (after conditioning).[10] There is very little LD between these two SNPs and conditioning on both showed that there was no residual signal in the region beyond these.[10]

Unfortunately rs12048904 was not included on the ImmunoChip, however a near perfect proxy rs12027668 was included. In the ImmunoChip data the evidence for association in the region was less pronounced (regression to the mean) and in these data the most associated (lead) SNP was found to be rs7552544. This SNP is in modest LD with both rs11581062 and rs12027668 (the proxy for rs12048904) and in both cases the risk alleles are positively correlated, indicating that association at rs7552544 is driven by both signals. Conditioning the ImmunoChip data on rs7552544 we found no residual signal at the most strongly correlated SNP, rs12027668 ($p = 9.1 \times 10^{-1}$), but nominally significant evidence of association at rs11581062 ($p = 1.2 \times 10^{-2}$) was apparent, as would be expected. Based on the ImmunoChip data alone one would likely have concluded that association in this region was modest and driven by a single variant (either rs7552544 or a variant in LD), however by chance these signals were stronger in the GWAS data set[10] where it was clear that association in the region is driven by two independent signals both in modest LD with rs7552544. In short in the ImmunoChip data it was not possible to resolve the modest apparent signal to the two independent signals underlying it. Supplementary Table 4 summarizes the association results for these SNPs and the pattern of LD between them.

Chromosome 10p15 (in the region of *IL2RA*)

Our 2007 GWAS[2] and related follow up efforts[6] identified rs2104286 as the most strongly associated SNP in this region. Although this SNP was not included in our most recent GWAS[10] the association was still identified, with the most associated (lead) SNP in the latest GWAS being rs3118470. Conditioning on this lead SNP revealed association with a second SNP rs7090512 (again both of these signals replicated and gave genomewide significant association in the combined analysis, after conditioning).[10] However the substantial LD between these two SNPs meant that effect on risk was best related to haplotypes of these two SNPs rather than to individual alleles (see supplementary file from the original GWAS[10]).

ImmunoChip included rs2104286 and rs7090512 but not rs3118470 (which failed QC). A reasonable proxy for rs3118470 was included on the ImmunoChip (rs4147359). Of note all these SNPs are in LD with each other. The most associated SNP in the ImmunoChip (rs2104286) seems to account for the whole signal in this region, with no evidence for any association left at rs4147359 (the proxy for rs3118470) or rs7090512 after conditioning. Supplementary Table 5 summarizes the association results for these SNPs and the pattern of LD between them.

Chromosome 3q13 (in the region of *CD86* and *SLC15A2*)

This is the most complex of the regions containing second signals identified in our most recent GWAS.[10] Conditioning on the most strongly associated SNP from the region

(rs9282641) not only revealed a second associated SNP (rs4285028) but after conditioning on both of these a third signal was also apparent (rs4308217). All three signals replicated and were genomewide significant in the combined analysis, however, as with the *IL2RA* region, the LD between two of the SNPs (rs9282641, the primary signal and rs4308217 the tertiary signal) meant that risk was best assigned to haplotypes of these two SNPs rather than individual alleles. Of these three SNPs only rs9282641 was included on ImmunoChip, however rs12695416 (a proxy for rs4285028) and rs2255214 (a proxy for rs4308217) were also included.

In the ImmunoChip data the most strongly associated SNP was rs1920296 ($p=6.8x10^{-15}$) which is a near perfect proxy for rs12695416 ($r^2=0.961$, D'=1.0) and is thus unsurprisingly in LD with the secondary signal from the GWAS (rs4285028; $r^2=0.373$, D'=0.80), and not with either of the other two SNPs identified in the GWAS ($r^2<0.04$ for both); what had been the secondary signal in the GWAS is actually the strongest signal in the ImmunoChip data set.

After conditioning on rs1920296, the most strongly associated SNP in the ImmunoChip data is rs2255214 (a proxy for the GWAS tertiary SNP rs4308217) although the original GWAS primary SNP (rs9282641) still yielded significant evidence for association. Conditioning on both rs1920296 and rs2255214 confirms some evidence for a residual signal from rs9282641. The analysis of this region is limited as it was not selected for fine mapping. The residual evidence for a haplotype effect suggests that an untyped variant from the region might still be responsible for the primary and tertiary signals identified in the GWAS. Supplementary Table 6 summarizes the association results for these SNPs and the pattern of LD between them.

Materials

Samples

All cases included in this study were diagnosed by Neurologists familiar with multiple sclerosis in accordance with recognised diagnostic criteria that employ a combination of clinical and laboratory-based para-clinical information.[15-17] The clinical characteristics of the patients are typical and vary only modestly between groups according to local interests and ascertainment strategies as outlined below. Disease severity was measured using the Expanded Disability Status Score (EDSS)[18] and its dependent derivative the Multiple Sclerosis Severity Score (MSSS),[19] while clinical features related to course, relapse and progression were defined in accordance with consensus criteria.[16,20-22] For cases analysed in the discovery phase, data regarding age at examination (AAE), age at onset (AAO), EDSS, MSSS and clinical course (primary progressive or bout onset) were available from 12327 (81%), 10795 (71%), 9631 (64%), 7934 (50%) and 12274 (81%) individuals respectively. See Supplementary Table 9 for clinical and demographic features of the cases.

As anticipated the index cases from the trio families were generally younger than patients in the case collections (reflecting the requirement for both parents to also be able to donate a DNA sample).

Local ascertainment

The ascertainment procedures for cases and controls from each population are described below; the non-European individuals genotyped as part of this project will be reported and described elsewhere. All cases and controls involved in this study gave valid informed consent in accordance with approval from the relevant local ethical committees or institutional review boards (IRBs).

Australia and New Zealand

All cases were self-identified volunteers recruited at centres located in Adelaide, Brisbane, Gold Coast, Hobart, Melbourne, Newcastle, Perth and Sydney and in New Zealand. All were confirmed by Neurologists.

Belgium

Samples were collected under coordination of the Neurology Department of the University Hospitals Leuven amongst out-patients and hospitalized patients with definite MS attending the Neurology Department of the University Hospitals Leuven or the "National MS Center" in Melsbroek (cases) and spouses of patients attending the Neurology Department of the University Hospitals Leuven (controls). Both centres are located 28 km apart in the centre of Belgium and recruited mainly amongst patients from the northern Dutch-speaking part of Belgium. Participation rate of patients attending these clinics is virtually 100%. At both centres, the majority of patients are followed-up longitudinally by neurologists specialized in MS with at least yearly visits. Approximately 40% of patients are being treated with an immunomodulatory therapy.

Denmark

Patients were recruited between 1996-2009 by neurologists at multiple sclerosis centers from across the whole of Denmark, although the majority of patients originate from the Copenhagen area. This clinic based approach means that the proportion of patients with relapse remitting multiple sclerosis (RRMS) is higher than is seen in the general population. Controls originate from staff members (10%) and healthy donor controls from the University Hospital Rigshospitalet in Copenhagen.

Finland

Cases were recruited from seven centres (Helsinki University Central Hospital, Tampere University Hospital, Kuopio University hospital, Oulu University Hospital, Seinäjoki Central Hospital, Satakunta Central Hospital and Rovaniemi Central Hospital) and thus come from many different regions of Finland. All were identified in hospital clinics by experienced neurologists. Part of the families have been recruited before 1998 as part of an effort to collect either multiplex families (at least two cases in a family) or trio/nuclear families (an affected individual with both parents and if not available with one parent and siblings), and the rest were recruited between 2000 and 2006 as trio/nuclear families by neurologists of the MGEN consortium.
Data from the Dietary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome study (DILGOM) subset of the national FINRISK 2007 Study were used as controls.[23] All individuals range in age from 25-74 years and were collected from five large study areas in Finland.

France

The French MS Genetics group (REFGENSEP) has been prospectively collecting samples with twenty-three centers covering France. Some patients volunteered in response to advertising campaigns through patient associations. Two annual meetings are organized for the national network REFGENSEP physicians in order to review procedures and diagnosis guidelines. Clinical information, informed consent and a blood sample were collected under the supervision of a trained physician. Control samples have been collected by the Etablissement Français du Sang (EFS) and come from all over France.

Germany

Three German centers contributed cases and controls:

a) Hamburg

Samples were collected from patients regularly seen in the MS outpatient clinic and day hospital of the Institute for Neuroimmunology and Clinical MS Research. The vast majority of these cases were derived from the local, i.e. Metropolitan region Hamburg, capture area. Hence, patient origin is largely from Northern Germany with the vast majority being of Caucasian descent. Control samples were also largely derived from this area, but since they were Institute and University employees, their geographic origin is more diverse, i.e. mostly from entire Germany, than the patients.

b) Munich

Samples from the TU in Munich can be stratified in three cohorts of cases and one cohort of controls. The first cohort comprises patients with multiple sclerosis from central Germany. The second cohort was recruited from across multiple sites in Germany and includes individuals with multiple sclerosis being treated with interferon-beta for at least 6 months. The third cohort of patients with multiple sclerosis was recruited in South-Eastern Germany. A cohort of controls was contributed from the population based study KORA (Cooperative Health Research in the Augsburg Region).

c) University Mainz

Samples were collected from outpatient clinics. Patients originate from different regions in Germany with the vast majority being of self-reported Caucasian descent. Collection was performed in Outpatient Clinics. Control samples were derived from university staff members of Caucasian descent.

Italy

Two Italian centers contributed cases and controls:

a) Piedmont

Patients were collected from continental Italy (excluding Sardinia) as part of the PROGEMUS (PROgnostic GEnetic factors in MUltiple Sclerosis) project, 87% of cases were collected in North-West Italy (Novara, Torino, Milano, Pavia) and 12% in Central Italy (Rome). These patients were all recruited from hospital based clinics; mean participation rate was approximately 60% (range 20%-90%).

Controls included Italian individuals (medical students, university and hospital staff, blood donors) matched for regional origin with the MS patients.

b) San Raffaele, Milan
Patients were recruited between 2000 and 2010 by MS neurologists at the Ospedale San Raffaele (OSR) in Milan. They are outpatients and hospitalized patients from continental Italy, with a higher proportion from northern Italy. Controls originate from spouses, friends or caregivers of affected patients and from staff members of the Ospedale San Raffaele.

Norway
The Norwegian MS samples were derived from two sources; the Oslo MS DNA biobank and the Norwegian Multiple Sclerosis Registry and Biobank held in Bergen. In the Oslo MS DNA biobank the majority of patients are recruited by the neurologists at Oslo University Hospital, Ullevål with the remainder coming from local MS Societies serving the suburban Oslo areas. Samples in the Norwegian Multiple Sclerosis Registry and Biobank were recruited from all other parts of Norway. This collection started in 2007, and currently includes approximately 1/5 of the prevalent MS patients in Norway. The Norwegian healthy controls were collected by the Norwegian Bone Marrow Donor Registry (http://www.oslo-universitetssykehus.no/omoss/avdelinger/benmargsgiverregisteret/).

Sweden
The Swedish cases and controls were recruited from four different investigations. Two concern the role of genetic and environmental risk factors in the development of multiple sclerosis: 1. The Epidemiological Investigation of Multiple Sclerosis (EIMS)[24] which is a nationwide population-based case-control study of incident multiple sclerosis cases (1098 cases and 1134 controls). 2. Genetic Environment study in Multiple Sclerosis (GEMS) which is a nation-wide population-based case-control study of prevalent cases of MS (1554 cases and 950 controls). A third investigation is a nation-wide study of natalizumab treatment in multiple sclerosis patients (IMSE) (963 cases).[25] A fourth study is collection of samples collected during routine neurological diagnostic work-up at Karolinska University Hosptial and Danderyd Hospital, Stockholm, Sweden (455 MS patients).[26] In addition samples from 721 blood donors from the Stockholm area with Scandinavian ancestry are included.[27] There was some overlap of patients between studies. In total 2839 independent cases and 2800 independent controls were included in the current investigation.

United Kingdom (UK)
The majority of UK cases were collected through a national recruitment project ("the genetic analysis of multiple sclerosis") coordinated by the Department of Clinical Neurosciences at the University of Cambridge and involving additional recruitment centres based in UK cities - Aberdeen, Birmingham, Bristol, Cardiff, Exeter, Hull, Ipswich, Leicester, London, Manchester, Newcastle, Norwich, Nottingham, Oxford, Peterborough, Preston, Plymouth, Poole, Rotherham, Sheffield, Southampton and Stoke. These cases were supplemented by additional cases recruited as part of local natural

history studies in the South West of England[28] and South Wales,[29] together with samples recruited as part of the Northern Isles Multiple Sclerosis (NIMS) study [30] and samples from the UK multiple sclerosis tissue bank (http://www.ukmstissuebank.imperial.ac.uk/). For a proportion of cases DNA was also collected from both parents to establish trio families (an affected individual and both parents); ultimately genotypes passing quality control in all three individuals were available from 621 trio families.
UK controls were obtained from three sources: the National Blood Service (anonymised DNA samples stored in the UK blood transfusions services repository that was originally established to support the activities of the Wellcome Trust Case Control Consortium), the 1958 Birth Cohort (DNA samples obtained from subjects involved with the National Child Development Study, an epidemiological survey based on all individuals born in England, Wales and Scotland during one week in 1958, see www.b58cgene.sgul.ac.uk/followup.php) and the NIMS study (we included these controls to more accurately match the cases included from the Scottish Isles).

United States of America (USA)
Four centers in the USA contributed cases and controls:

a) Brigham & Women's Hospital (BWH), Boston MA
Study participants were recruited through the Partners MS Center in Boston, MA. All samples were collected at the MS Center and processed on site to extract DNA. Healthy control subjects were contributed by the BWH PhenoGenetic project, a collection of subjects 18-50 years of age living in the Boston metropolitan area who are self-reported to be healthy.

b) University of Miami (UM)
Study participants were recruited using multiple ascertainment approaches. However, the majority of participants were enrolled through the University of Miami Health System's designated MS Center of Excellence. Additional participants were recruited via MS community outreach events and support group meetings. Data on the clinical characteristics and clinical history of MS cases were obtained via medical chart review by a board-certified neurologist. Control participants included unrelated spouses and friends of affected individuals in addition to unaffected individuals in the general population. These control subjects were recruited via the same ascertainment mechanisms.

c) University of California San Francisco (UCSF).
Study participants were recruited from the UCSF MS clinic and from other collaborating sites across the United States using common inclusion and exclusion criteria. Phlebotomy was performed at the individual's preferred clinic, and blood samples were shipped to the UCSF laboratory by overnight courier. The dataset studied here is comprised of two groups, multicase families, in which at least one first-degree relative of the affected proband also had clinically definite MS, and sporadic cases, in which the affected individual reported no known history of MS in any family member. MS phenotypes were confirmed by systematic chart review. All known ancestors were of European descent. Controls were also of European ancestry and consist primarily of spouses and friends of

MS patients who reported no known history of chronic diseases, including in first-degree relatives.

d) Vanderbilt University (Nashville)
All samples were obtained from the Vanderbilt DNA Biorepository (BioVU). This repository connects a de-identified version of Vanderbilt's electronic medical records to DNA samples extracted from waste blood obtained from the phlebotomy labs at Vanderbilt. Thus all samples were obtained through the Vanderbilt Hospitals and Clinics and represent a primary catchment area of Middle Tennessee, U.S.A. Cases were defined using algorithms focused on ICD-9 billing codes, prescribed MS treatments, and keywords located in the text. A manual review of 50 cases indicates a positive predictive value of 98%. Controls were defined as having no ICD-9 code of 340, no mention of multiple sclerosis, no use of medications commonly prescribed for multiple sclerosis, and no mention of any other autoimmune disease.

DNA

Standard methods were used to extract DNA and each contributing centre quantified and normalised their samples locally. Cases and controls from France and the USA were genotyped at local centres as were half of the German (KORA) controls. Samples from Italy and Sardinia, and one third of the samples from Sweden, were genotyped at the Miami University, John P. Hussman Institute for Human Genomics. All other samples (approximately two thirds of the total) were genotyped at the Wellcome Trust Sanger Institute (WTSI). Sample quality control and renormalisation was performed at the respective centres prior to genotyping.

Genotyping

We genotyped a total of 45885 samples, with typing being performed in Boston (1824), France (1028), Germany (997), Miami (5383), Virginia (5416), Vanderbilt (3817) and the Wellcome Trust Sanger Institute (WTSI, 27420). All typing was performed according to the manufactures standard specifications. Raw data from all sites was transferred to the WTSI and was called in three batches using OptiCall (http://www.sanger.ac.uk/resources/software/opticall/).[31] The Wellcome Trust Case Control Consortium common control samples from the 1958 birth cohort (6894), the UK National blood transfusion service (3057) and the HapMap project (48) were called in one batch (these control data were used by other groups working with the ImmunoChip on other diseases). All other multiple sclerosis related data generated at the WTSI were called together as a second batch (22837). All multiple sclerosis related data that was generated outside the WTSI was called as a separate third batch (13049). In total 4149 samples failed QC (see below). The remaining samples included 1230 African Americans, 244 Hispanics, 321 South Asians, 742 individuals from 70 multiplex families, 1899 individuals from 633 trio families (an affected individual and both parents), 17445 unrelated cases and 19855 unrelated controls. We excluded the 1795 non-Europeans from this analysis (these data will be reported elsewhere in an independent analysis).

Sample quality control (QC)

Sample QC was performed in a hierarchal fashion. For each quality test applied, the number of samples excluded is shown, with the breakdown of this number in terms of cases and controls shown in parentheses - number excluded (cases:controls).

Samples were excluded if

1) The observed gender based on sex chromosome markers disagreed with the reported gender - 171 (51:120). This assessment was based on calls generated using Illuminus[32] since at the time of analysis OptiCall[31] was not configured to process sex chromosome markers.
2) Call Rate < 98% across all 192402 markers that generated data - 192 (72:120). Note sample QC was done prior to any SNP QC.
3) Autosomal heterozygosity more than 3 standard deviations from the mean value - 230 (119:111)
4) Ambiguity or inconsistency in the sequenom finger print ID - 924 (189:735). In most cases we were able to establish that such errors resulted from technical issues such as plate swaps, faulty chips or sample handling errors and we were therefore able to attempt retyping of these samples.
5) Excessive Identity By Decent (IBD) - 991 (441:544 and 6 unknown). Using the IBD command in PLINK we made all pairwise sample comparisons and excluded the sample with the lowest call rate from non-family pairs with PI_HAT $\geq$ 0.25; PI_HAT being defined as P(IBD=2)+0.5*P(IBD=1). In total 686 samples were judged to have been typed in duplicate (i.e. had PI_HAT > 0.90). Some families had to be excluded because they were related to other families or were found to be unrelated to each other. Within the families surviving QC all pairwise comparisons had PI_HAT between 0.25 and 0.60.
6) Eigenstrat outliers - 1330 (721:609). Principal components (PC) were generated based on each of the 11 populations considered (see below) and individuals were excluded if they were more than 6 standard deviations from the mean on any of the first ten PCs within their respective population. All individuals were also projected onto HapMap PCs and excluded if they were outliers with respect to the European group. Some families had to be excluded because they were non white.
7) Excessive Mendelian Errors - 51 (18:33). On testing the trio and multiplex families we found and excluded 17 trio families with > 5000 Mendelian errors, there were no multiplex families showing this level of inconsistency.
8) Other - 165 (14:151). In some families the parents (one or both) had to be excluded because other key individuals in that family failed QC (e.g. the index case in a trio family). Some samples were removed because they withdrew from the study.

A population specific break down of these exclusions is provided in Supplementary Table 10.

Shared external controls

The International Inflammatory Bowel Disease Genetics Consortium (IIBDGC, http://www.ibdgenetics.org) provided us with ImmunoChip genotypes from 20337

control individuals, of which 9799 were found to overlap with control individuals already included in our study. A further 433 of the IIBDGC samples failed our QC so that in total we included data from 10102 of the IIBDGC samples, increasing the number of controls available for all populations except Finland, France, Norway and the UK.

Overlap with existing GWAS

Through Identity By Descent calculations, it was determined that 8813 samples genotyped on ImmunoChip (2947 cases and 5866 controls) overlapped with samples which were part of the previous GWAS efforts [10,11]. These samples were removed from the discovery phase and included in the replication phase to enable an independent and robust replication using data from an already completed meta-analysis of all previous GWAS (14802 cases and 26703 controls).

SNP QC

The QC for SNPs was applied in each of the 11 population strata independently using just those samples that passed QC. SNP QC was performed in a hierarchal fashion. In each stratum, SNPs were excluded if they

1) Had a call rate < 98%

2) Showed significant evidence of deviation from Hardy-Weinberg Equilibrium, with $p < 1.0 \times 10^{-5}$

3) Showed evidence of differential missingness between cases and controls with $p < 1.0 \times 10^{-3}$

4) Were monomorphic

Numbers excluded under each criteria are shown in Supplementary Table 11.

Amongst the full set of 192402 SNPs there were 839 duplicates and 13952 that gave more than 1 Mendelian error across our 703 families (633 trio families and 70 multiplex families). These SNPs were excluded in all populations. Since the German controls were typed in two parts (KORAEX V and KORAF4) at different centres we compared these data. Inspection of the Quantile-Quantile (QQ) plots revealed 18 SNPs that showed notable deviation from the expected null distribution. Thus, we excluded these SNPs in the German analysis. The same process was repeated for Sweden and the UK (two other sites where controls were processed in two parts at different centres), this led to the exclusion of 8 and 6 SNPs in Sweden and the UK respectively.

The external IIBDGC control data included genotypes from 8076 UK control subjects that had also been called as part of our effort. Comparing these duplicate calls identified 776 SNPs with a significant difference ($p < 1.0 \times 10^{-3}$), these were excluded in each of the populations where the IIBDGC external controls were employed. The IIBDGC and multiple sclerosis controls were compared within each population where both were available. This identified a total of 678 SNPs showing a significant difference ($p < 1.0 \times 10^{-5}$) in at least one population. These were excluded in all populations which included IIBDGC controls (i.e. all except Finland, France, Norway and the UK). After a preliminary case control analysis potentially associated markers were identified and their cluster plots were visually inspected using Evoker (http://www.sanger.ac.uk/resources/software/evoker/).[33] In total we checked the cluster plots from 4896 SNPs and rejected 197 as inadequate, these 197 SNPs were excluded

from subsequent analysis, see Supplementary Figure 98 for examples of rejected SNPs. After all these exclusions there were a total of 165892 SNPs available in at least one population. Markers available in only one population were excluded from analysis, resulting in 161328 SNPs analysed.

Principal Component Analysis

Principal Components (PCs) were calculated in each of the 11 stratum independently, using a core set of 21468 SNPs. Only samples which passed all quality control were included. For this purpose of accounting for population substructure in association analysis, no HapMap samples were included as they were for outlier detection.
These 21468 remained after excluding SNPs

1) That did not pass QC in any 1 of the 11 data sets
2) From the MHC regions (all markers on chromosome 6 from 26 to 36 MB)
3) From other regions of extended LD such as chromosome 8 from 6 to 16 MB and chromosome 17 from 40 to 45 MB
4) With a minor allele frequency (MAF) less than 0.01
5) From LD pruning so that no marker had pairwise r^2 >0.1 with any marker within a 100 SNP window.

By way of an example Supplementary Figure 97 shows the UK cases and controls plotted according to the first two PCs.
After adjusting to a standard sample size[34] only Denmark and Norway showed evidence of any substantial genomic inflation (see Supplementary Table 12). After including the first 5 PCs as covariates, only minimal evidence for genomic inflation in Denmark remained (final column in Supplementary Table 12).

Duplicate Sample and SNP Concordance

There were 9312 unique duplicate sample pairs (i.e. had PI_HAT > 0.90), where both samples in the duplicate pair passed all other sample quality control (i.e. one sample in each pair was dropped from the final analysis only because of duplication). We checked the concordance of each of these duplicate pairs in the 161311 clean SNPs used in the final analysis. For each pair, we checked the concordance only in the SNPs for which the sample included in the final analysis was analysed. Of the 9312 duplicate pairs, 462 were pairs containing two IMSGC samples and the other 8850 were pairs containing one IMSGC sample and one IIBDGC sample. The concordance rate for the 9312 unique sample pairs was 0.99892. In the 462 pairs where both were from the IMSGC, the concordance rate was 0.99888, and in the 8850 pairs where one sample was from IIBDGC the concordance rate was 0.99953. Restricting the concordance check to the now identified 109 multiple sclerosis risk variants genotyped on the ImmunoChip, the concordance rates were 0.99997, 0.99946, and 1.0 for the 9312, 462, and 8850 pairs respectively.

There were 833 unique duplicate SNP pairs where both SNPs in the duplicate pair passed all other SNP quality control (i.e. one SNP in each pair was dropped from the final analysis only because of duplication). We checked the concordance of each of these duplicate pairs in the 38589 clean samples used in the final analysis. For each pair, we

checked the concordance only in the samples for which the SNP included in the final analysis was analysed. The concordance rate for the 833 unique SNP pairs was 0.99951.

Additional Analysis

Secondary phenotypes

In addition to the primary analysis of susceptibility we also performed an analysis of the ImmunoChip data from the cases with respect to severity (as measured in terms of the MSSS[19]) In this severity (MSSS) based analysis, data from each population were tested separately using the first five principal components together with age at onset and gender as covariates, and then corrected for genomic inflation before results were combined in meta-analysis across the 11 populations.

No SNP reached genomewide significance in this secondary phenotype analysis. The most associated SNP in the MSSS ImmunoChip analysis was rs4092077 from an intergenic region on chromosome 4q35. Supplementary Figure 99 shows the QQ plot for the MSSS ImmunoChip analysis.

Trio families

In addition to the main case control samples we also typed the ImmunoChip in a set of trio families (an affected individual and both parents). After QC this effort provided data from all three individuals in 633 trio families (621 from the UK and 12 from the USA). Transmission disequilibrium testing (TDT) in these families showed over transmission of the risk allele apparent in the case control analysis for all but 16 of the 97 independent genomewide significant SNPs identified. None of the 16 discordant SNPs showed statistically significant evidence of deviation from the null while 25 of the concordant SNP showed a significant difference (one sided p-value 5.0×10^{-2}). The probability of seeing 81/97 concordant over transmission by chance alone is 6.2×10^{-12}. Supplementary Figure 100 shows the QQ plot for the TDT analysis.

Heritability Explained

In order to calculate the genetic variance currently explained in multiple sclerosis, we used a logistic regression model in the discovery phase data. Using a joint analysis of all discovery phase data for association with multiple sclerosis, we fit a null model with the first 5 principal components and indicator variables for the 11 country-level strata as covariates. We then fit an alternative model which additionally included the 109 non-MHC susceptibility alleles that were included on ImmunoChip (see Supplementary Table 8) as well as tagging SNPs available in all 11 strata (rs2523822 for HLA-A*02:01, rs3135388 for DRB1*15:01, rs1265754 for DRB1*03:01, and rs3763308 for DRB1*13:03) for the four MHC risk alleles presented in the GWAS.[10] If the 109 markers were not available in all 11 strata, we chose a tag with $r^2 > 0.4$ This was necessary for only 8 of 109 SNPs. In this way, the variance estimate is likely conservative. We estimate that the 110 non-MHC established risk variants explain 18%

of the variance in the data; 27% including the already identified MHC effects. Variance was calculated using the Nagelkerke R2 in R version 2.13.

In the 2011 GWAS[10] we calculated that collectively the four MHC and 57 non-MHC susceptibility alleles identified in that study accounted for 25% of the sibling recurrence risk (λ_s) observed epidemiologically; assuming a total λ_s of 6.3.[35] In addition we also used a liability threshold model and estimated that these same risk alleles explain 17% of the genetic variance.

Using the same methods as described in the supplementary information file from the 2011 GWAS[10] we repeated these calculations using the summary statistics from the109 non-MHC susceptibility alleles that were included on ImmunoChip. These variants explain a λs of 1.271 and 12.8% of the genetic variance in liability. Combined with the effects attributable to the four MHC risk alleles[10] we now explain a total of 28% of the sibling recurrence risk and 23% of the genetic variance in liability. Given the geometric relationship between recurrence risk and relatedness apparent in multiple sclerosis[36,37] it is likely that a substantial fraction of the apparent heritability of the disease is phantom,[38] (i.e. attributable to interactions rather than the additive marginal effects identified in GWAS) and therefore that the proportion of narrow sense heritability explained is likely to be substantially higher.

Acknowledgements

We thank the patients, families and healthy control individuals that participated in this study. We also thank the nurses and referring physicians. This project was supported by the National Institutes of Health (NS049477, NS26799, R01NS032830, RC2NS070340, R01NS067305, RC2GM093080), the Wellcome Trust (as part of the Wellcome Trust Case Control Consortium 2 project - 085475/B/08/Z, 085475/Z/08/Z, 084702/Z/08/Z and 098051), the Multiple Sclerosis Society of Great Britain and Northern Ireland (857/07, 861/07, 862/07, 894/08, 898/08 and 955/11), the UK Medical Research Council (G0700061), the US National Multiple Sclerosis Society (grants RG 4198-A-1 and 4680-A-1, postdoctoral fellowship FG 1938-A-1 and Harry Weaver Neuroscience Scholars JF2138A1 and JF-2137A4), the South Florida chapter of the National Multiple Sclerosis Society, the Cambridge NIHR Biomedical Research Centre, DeNDRon North West, the Bibbi and Niels Jensens Foundation, the Swedish Brain Foundation, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Foundation for Strategic Research, the Stockholm County Council (project 592229), the Strategic Cardiovascular and Diabetes Programmes of Karolinska Institutet and Stockholm County Council, the Swedish Council for Working life and Social Research, the Institut National de la Santé et de la Recherche Médicale (INSERM), the Fondation d'Aide pour la Recherche sur la Sclérose En Plaques (ARSEP), the Association Française contre les Myopathies (AFM) and GIS-IBISA, the German Ministry for Education and Research (BMBF), the German Competence Network Multiple Sclerosis (KKNMS, Control-MS, 01GI0917), the Deutsche Forschungsgemeinschaft (SyNergy Cluster, JE 530/1-1), Munich Biotec Cluster M4, The Dutch MS Research Foundation, the Fidelity Biosciences Research Initiative, the Research Foundation Flanders (FWO-Vlaanderen), the Research Fund KU Leuven (OT/11/087), the KU Leuven Biogen Idec Chair of Translational Research in Multiple Sclerosis, the Belgian Charcot Foundation, the Gemeinnützige Hertie Stiftung, the Clinical Research Priority Program-MS (CRPPMS) of the University Zurich, the Danish Multiple Sclerosis Society, the Danish Council for Strategic Research, the Center of Excellence for Disease Genetics of the Academy of Finland, the Sigrid Juselius Foundation, the Helsinki University Central Hospital Research Foundation, the Italian Foundation for Multiple Sclerosis (FISM grants "Progetto Speciale Immunochip", 2011/R/14), Fondazione Cariplo (grant 2010-0728), the Italian Ministry of University and Research (MIUR, PRIN08), the CRT Foundation of Turin (Italy), the Italian Ministry of Health (grant Giovani Ricercatori 2007, D.lgs 502/92), the Italian Foundation for Multiple Sclerosis, the Italian Institute of Experimental Neurology (INSPE), the MS association of Oslo, the Norwegian Research Council (143153 and 143410/HFH), the South Eastern Norwegian Health Authorities (51852/ILM), the Australian National Health and Medical Research Council (633275) and Northern California Kaiser Permanente members. This research was funded in part by NIH/NIAID R01AI076544, NIH/NIEHS R01ES017080, and NIH/NINDS R01NS049510. We also remember the contributions and tireless efforts of our departed friend and colleague Prof Leena Peltonen.

We acknowledge use of samples from the British 1958 Birth Cohort DNA collection (funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant

068545/Z/02), the UK National Blood Service controls (funded by the Wellcome Trust), the Vanderbilt University Medical Center's BioVU DNA Resources Core (supported by institutional funding and by the Vanderbilt CTSA grant 1UL1RR024975-01 from NCRR/NIH), the Centre de Ressources Biologiques du Réseau Français d'Etude Génétique de la Sclérose en Plaques (CRB-REFGENSEP), the Norwegian Bone Marrow Registry controls, the Norwegian Multiple Sclerosis Registry and Biobank, the North American Research Committee on MS (NARCOMS) Registry (supported by the Consortium of Multiple Sclerosis Centers, CMSC) and the Brigham and Womens Hospital PhenoGenetic Project. This study makes use of data generated by the DILGOM project funded by the Academy of Finland (grants 136895, 263836, 118065).
The GWAS made use of external control data from the popgen biobank (www.popgen.de, supported by the German Ministry of Education and Research (BMBF) through the National Genome Research Network (NGFN) and received infrastructure support through the DFG excellence cluster "Inflammation at Interfaces"), the Swedish Breast Cancer study (funded by the Agency for Science & Technology and Research of Singapore (A*STAR), the Susan G Komen Breast Cancer Foundation, and the National Institute of Health - R01 CA 104021), HYPERGENES (a European Commission project, HEALTH-F4-2007-201550), the Children's Hospital of Philadelphia (CHOP, funded by an Institutional Development Award to the Center for Applied Genomics from the Children's Hospital of Philadelphia), the Swedish CAD study (funded by the Knut and Alice Wallenberg Foundation, the Swedish Research Council - 8691 and the Stockholm County Council - 562183) and the collaboration of the BRC-REFGENSEP, Pitié-Salpêtrière Centre d"Investigation Clinique (CIC) and Généthon.
We thank the Wellcome Trust Sanger Institute Genotyping Facility (particularly Emma Gray, Sue Bumpstead and Doug Simpkin), the Biorepository and the Center for Genome Technology within the University of Miami John P. Hussman Institute for Human Genomics (specifically Sandra West and Patrice Whitehead), the technical staff, medical assistants, nursing and medical staff in the Center for Applied Genomics, Children's Hospital of Philadelphia, the Institut du Cerveau et de la Moelle épinière (ICM), CIC Pitié-Salpêtrière, Généthon and REFGENSEP's members, colleagues from Finland (Tuula Pirttilä, Veikko Salomaa, Eveliina Jakkula), the Norwegian Multiple Sclerosis Registry and Biobank (Jan Aarseth), Italy (Bruno Colombo, Lucia Moiola, Paolo Rossi, Marta Radaelli and Cinzia Pozzoni), the SNP Technology Platform in Uppsala (www.genotyping.se) and UCSF (Stacy Caillier, H. Mousavi, R. O'Shea, and A. Santaniello) for their help and support.

Membership of Wellcome Trust Case Control Consortium 2

Management Committee
Peter Donnelly (Chair)[1,2], Ines Barroso (Deputy Chair)[3], Jenefer M Blackwell[4, 5], Elvira Bramon[6] , Matthew A Brown[7] , Juan P Casas[8] , Aiden Corvin[9], Panos Deloukas[3], Audrey Duncanson[10], Janusz Jankowski[11], Hugh S Markus[12], Christopher G Mathew[13], Colin NA Palmer[14], Robert Plomin[15], Anna Rautanen[1], Stephen J Sawcer[16], Richard C Trembath[13], Ananth C Viswanathan[17], Nicholas W Wood[18]
Data and Analysis Group
Chris C A Spencer[1], Gavin Band[1], Céline Bellenguez[1], Colin Freeman[1], Garrett Hellenthal[1], Eleni Giannoulatou[1], Matti Pirinen[1], Richard Pearson[1], Amy Strange[1], Zhan Su[1], Damjan Vukcevic[1], Peter Donnelly[1,2]
DNA, Genotyping, Data QC and Informatics Group
Cordelia Langford[3], Sarah E Hunt[3], Sarah Edkins[3], Rhian Gwilliam[3], Hannah Blackburn[3], Suzannah J Bumpstead[3], Serge Dronov[3], Matthew Gillman[3], Emma Gray[3], Naomi Hammond[3], Alagurevathi Jayakumar[3], Owen T McCann[3], Jennifer Liddle[3], Simon C Potter[3], Radhi Ravindrarajah[3], Michelle Ricketts[3], Matthew Waller[3], Paul Weston[3], Sara Widaa[3], Pamela Whittaker[3], Ines Barroso[3], Panos Deloukas[3].
Publications Committee
Christopher G Mathew (Chair)[13], Jenefer M Blackwell[4,5], Matthew A Brown[7], Aiden Corvin[9], Chris C A Spencer[1]

1 Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; 2 Dept Statistics, University of Oxford, Oxford OX1 3TG, UK; 3 Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; 4 Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, 100 Roberts Road, Subiaco, Western Australia 6008; 5 Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge CB2 0XY, UK; 6 Department of Psychosis Studies, NIHR Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King's College London and The South London and Maudsley NHS Foundation Trust, Denmark Hill, London SE5 8AF, UK; 7 University of Queensland Diamantina Institute, Brisbane, Queensland, Australia; 8 Dept Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT and Dept Epidemiology and Public Health, University College London WC1E 6BT, UK; 9 Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Eire; 10 Molecular and Physiological Sciences, The Wellcome Trust, London NW1 2BE; 11 Department of Oncology, Old Road Campus, University of Oxford, Oxford OX3 7DQ, UK , Digestive Diseases Centre, Leicester Royal Infirmary, Leicester LE7 7HH, UK and Centre for Digestive Diseases, Queen Mary University of London, London E1 2AD, UK; 12 Clinical Neurosciences, St George's University of London, London SW17 0RE; 13 King's College London Dept Medical and Molecular Genetics, King's Health Partners, Guy's Hospital, London SE1 9RT, UK; 14 Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK; 15 King's College London Social, Genetic and Developmental

Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK; 16 University of Cambridge Dept Clinical Neurosciences, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK; 17 NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London EC1V 2PD, UK; 18 Dept Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK.

Membership of the WTCCC+ consortium

Jan Aerts[1], Tariq Ahmad[2], Hazel Arbury[1], Anthony Attwood[1,3,4], Adam Auton[5], Stephen G Ball[6], Anthony J Balmforth[6], Chris Barnes[1], Jeffrey C Barrett[1], Inês Barroso[1], Anne Barton[7], Amanda J Bennett[8], Sanjeev Bhaskar[1], Katarzyna Blaszczyk[9], John Bowes[7], Oliver J Brand[8,10], Peter S Braund[11], Francesca Bredin[12], Gerome Breen[13,14], Morris J Brown[15], Ian N Bruce[7], Jaswinder Bull[16], Oliver S Burren[17], John Burton[1], Jake Byrnes[18], Sian Caesar[19], Niall Cardin[5], Chris M Clee[1], Alison J Coffey[1], John MC Connell[20], Donald F Conrad[1], Jason D Cooper[17], Anna F Dominiczak[20], Kate Downes[17], Hazel E Drummond[21], Darshna Dudakia[16], Andrew Dunham[1], Bernadette Ebbs[16], Diana Eccles[22], Sarah Edkins[1], Cathryn Edwards[23], Anna Elliot[16], Paul Emery[24], David M Evans[25], Gareth Evans[26], Steve Eyre[7], Anne Farmer[14], I Nicol Ferrier[27], Edward Flynn[7], Alistair Forbes[28], Liz Forty[29], Jayne A Franklyn[10,30], Timothy M Frayling[2], Rachel M Freathy[2], Eleni Giannoulatou[5], Polly Gibbs[16], Paul Gilbert[7], Katherine Gordon-Smith[19,29], Emma Gray[1], Elaine Green[29], Chris J Groves[8], Detelina Grozeva[29], Rhian Gwilliam[1], Anita Hall[16], Naomi Hammond[1], Matt Hardy[17], Pile Harrison[31], Neelam Hassanali[8], Husam Hebaishi[1], Sarah Hines[16], Anne Hinks[7], Graham A Hitman[32], Lynne Hocking[33], Chris Holmes[5], Eleanor Howard[1], Philip Howard[34], Joanna MM Howson[17], Debbie Hughes[16], Sarah Hunt[1], John D Isaacs[35], Mahim Jain[18], Derek P Jewell[36], Toby Johnson[34], Jennifer D Jolley[3,4], Ian R Jones[29], Lisa A Jones[19], George Kirov[29], Cordelia F Langford[1], Hana Lango-Allen[2], G Mark Lathrop[37], James Lee[12], Kate L Lee[34], Charlie Lees[21], Kevin Lewis[1], Cecilia M Lindgren[8,18], Meeta Maisuria-Armer[17], Julian Maller[18], John Mansfield[38], Jonathan L Marchini[5], Paul Martin[7], Dunecan CO Massey[12], Wendy L McArdle[39], Peter McGuffin[14], Kirsten E McLay[1], Gil McVean[5,18], Alex Mentzer[40], Michael L Mimmack[1], Ann E Morgan[41], Andrew P Morris[18], Craig Mowat[42], Patricia B Munroe[34], Simon Myers[18], William Newman[26], Elaine R Nimmo[21], Michael C O'Donovan[29], Abiodun Onipinla[34], Nigel R Ovington[17], Michael J Owen[29], Kimmo Palin[1], Aarno Palotie[1], Kirstie Parnell[2], Richard Pearson[8], David Pernet[16], John RB Perry[2,18], Anne Phillips[42], Vincent Plagnol[17], Natalie J Prescott[9], Inga Prokopenko[8,18], Michael A Quail[1], Suzanne Rafelt[11], Nigel W Rayner[8,18], David M Reid[33], Anthony Renwick[16], Susan M Ring[39], Neil Robertson[8,18], Samuel Robson[1], Ellie Russell[29], David St Clair[13], Jennifer G Sambrook[3,4], Jeremy D Sanderson[40], Stephen J Sawcer[43], Helen Schuilenburg[17], Carol E Scott[1], Richard Scott[16], Sheila Seal[16], Sue Shaw-Hawkins[34], Beverley M Shields[2], Matthew J Simmonds[8,10], Debbie J Smyth[17], Elilan Somaskantharajah[1], Katarina Spanova[16], Sophia Steer[44], Jonathan Stephens[3,4], Helen E Stevens[17], Kathy Stirrups[1], Millicent A Stone[45,46], David P Strachan[47], Zhan Su[5], Deborah PM Symmons[7], John R Thompson[48], Wendy Thomson[7], Martin D Tobin[48], Mary E Travers[8], Clare Turnbull[16], Damjan Vukcevic[18], Louise V Wain[48], Mark Walker[49], Neil M Walker[17], Chris Wallace[17], Margaret Warren-Perry[16], Nicholas A Watkins[3,4], John Webster[50], Michael N Weedon[2], Anthony G Wilson[51], Matthew Woodburn[17], B Paul Wordsworth[52], Chris Yau[5], Allan H Young[27,53], Eleftheria Zeggini[1], Matthew A Brown[52,54], Paul R Burton[48], Mark J Caulfield[34], Alastair Compston[43], Martin Farrall[55], Stephen CL Gough[8,10,30], Alistair S Hall[6], Andrew T Hattersley[2,56], Adrian VS Hill[18], Christopher G Mathew[9], Marcus Pembrey[57], Jack Satsangi[21], Michael R Stratton[1,16], Jane Worthington[7], Matthew E Hurles[1], Audrey Duncanson[58], Willem H Ouwehand[1,3,4], Miles Parkes[12], Nazneen Rahman[16], John A Todd[17], Nilesh J Samani[11,59], Dominic P

Kwiatkowski[1,18], Mark I McCarthy[8,18,60], Nick Craddock[29], Panos Deloukas[1], Peter Donnelly[5,18].

[1] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA UK.
[2] Genetics of Complex Traits, Peninsula College of Medicine and Dentistry University of Exeter, EX1 2LU, UK.
[3] Department of Haematology, University of Cambridge, Long Road, Cambridge, CB2 0PT, UK.
[4] National Health Service Blood and Transplant, Cambridge Centre, Long Road, Cambridge CB2 0PT, UK.
[5] Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG, UK.
[6] Multidisciplinary Cardiovascular Research Centre (MCRC), Leeds Institute of Genetics, Health and Therapeutics (LIGHT), University of Leeds, Leeds, LS2 9JT, UK.
[7] arc Epidemiology Unit, Stopford Building, University of Manchester, Oxford Road, Manchester, M13 9PT, UK.
[8] Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, UK.
[9] Department of Medical and Molecular Genetics, King's College London School of Medicine, 8th Floor Guy's Tower, Guy's Hospital, London, SE1 9RT, UK.
[10] Centre for Endocrinology, Diabetes and Metabolism, Institute of Biomedical Research, University of Birmingham, Birmingham, B15 2TT, UK.
[11] Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Groby Road, Leicester LE3 9QP, UK.
[12] IBD Genetics Research Group, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK.
[13] University of Aberdeen, Institute of Medical Sciences, Foresterhill, Aberdeen AB25 2ZD, UK.
[14] SGDP, The Institute of Psychiatry, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK.
[15] Clinical Pharmacology Unit, University of Cambridge, Addenbrookes Hospital, Hills Road, Cambridge CB2 2QQ, UK.
[16] Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton SM2 5NG, UK.
[17] Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Cambridge CB2 0XY, UK.
[18] The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK.
[19] Department of Psychiatry, University of Birmingham, National Centre for Mental Health, 25 Vincent Drive, Birmingham, B15 2FG, UK.
[20] BHF Glasgow Cardiovascular Research Centre, University of Glasgow, 126 University Place, Glasgow, G12 8TA, UK.
[21] Gastrointestinal Unit, Division of Medical Sciences, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK.

[22] Academic Unit of Genetic Medicine, University of Southampton, Southampton, UK.
[23] Endoscopy Regional Training Unit, Torbay Hospital, Torbay TQ2 7AA, UK.
[24] Academic Unit of Musculoskeletal Disease, University of Leeds, Chapel Allerton Hospital, Leeds, West Yorkshire LS7 4SA, UK.
[25] MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Bristol, BS8 2BN, UK.
[26] Department of Medical Genetics, Manchester Academic Health Science Centre (MAHSC), University of Manchester, Manchester M13 0JH, UK.
[27] School of Neurology, Neurobiology and Psychiatry, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne, NE1 4LP, UK.
[28] Institute for Digestive Diseases, University College London Hospitals Trust, London NW1 2BU, UK.
[29] MRC Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, UK.
[30] University Hospital Birmingham NHS Foundation Trust, Birmingham, B15 2TT, UK.
[31] University of Oxford, Institute of Musculoskeletal Sciences, Botnar Research Centre, Oxford, OX3 7LD, UK.
[32] Centre for Diabetes and Metabolic Medicine, Barts and The London, Royal London Hospital, Whitechapel, London, E1 1BB, UK.
[33] Bone Research Group, Department of Medicine and Therapeutics, University of Aberdeen, Aberdeen, AB25 2ZD, UK.
[34] Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK.
[35] Institute of Cellular Medicine, Musculoskeletal Research Group, 4th Floor, Catherine Cookson Building, The Medical School, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK.
[36] Gastroenterology Unit, Radcliffe Infirmary, University of Oxford, Oxford, OX2 6HE, UK.
[37] Centre National de Genotypage, 2, Rue Gaston Cremieux, Evry, Paris 91057, France.
[38] Department of Gastroenterology & Hepatology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne NE1 4LP, UK.
[39] ALSPAC Laboratory, Department of Social Medicine, University of Bristol, BS8 2BN, UK.
[40] Division of Nutritional Sciences, King's College London School of Biomedical and Health Sciences, London SE1 9NH, UK.
[41] NIHR-Leeds Musculoskeletal Biomedical Research Unit, University of Leeds, Chapel Allerton Hospital, Leeds, West Yorkshire LS7 4SA, UK.
[42] Department of General Internal Medicine, Ninewells Hospital and Medical School, Ninewells Avenue, Dundee DD1 9SY, UK.
[43] Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 2QQ, UK.
[44] Clinical and Academic Rheumatology, Kings College Hosptal National Health Service Foundation Trust, Denmark Hill, London SE5 9RS, UK.
[45] University of Toronto, St. Michael's Hospital, 30 Bond Street, Toronto, Ontario M5B 1W8, Canada.

[46] University of Bath, Claverdon, Norwood House, Room 5.11a Bath Somerset BA2 7AY, UK.
[47] Division of Community Health Sciences, St George's, University of London, London SW17 0RE, UK.
[48] Departments of Health Sciences and Genetics, University of Leicester, 217 Adrian Building, University Road, Leicester, LE1 7RH, UK.
[49] Diabetes Research Group, School of Clinical Medical Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK.
[50] Medicine and Therapeutics, Aberdeen Royal Infirmary, Foresterhill, Aberdeen, Grampian AB9 2ZB, UK.
[51] School of Medicine and Biomedical Sciences, University of Sheffield, Sheffield, S10 2JF, UK.
[52] Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Nuffield Orthopaedic Centre, University of Oxford, Windmill Road, Headington, Oxford, OX3 7LD, UK.
[53] UBC Institute of Mental Health, 430-5950 University Boulevard Vancouver, British Columbia, V6T 1Z3, Canada.
[54] Diamantina Institute of Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Ipswich Road, Woolloongabba, Brisbane, Queensland, 4102, Australia.
[55] Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK.
[56] Genetics of Diabetes, Peninsula College of Medicine and Dentistry, University of Exeter, Barrack Road, Exeter, EX2 5DW, UK.
[57] Clinical and Molecular Genetics Unit, Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK.
[58] The Wellcome Trust, Gibbs Building, 215 Euston Road, London NW1 2BE, UK.
[59] Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, LE3 9QP, UK.
[60] Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, OX3 7LJ, UK.

International IBD Genetics Consortium (IIBDGC) contributing members

Murray Barclay[1], Laurent Peyrin-Biroulet[2], Mathias Chamaillard[3], Jean-Frederick Colombel[4], Mario Cottone[5], Anthony Croft[6], Renata D'Incà[7], Jonas Halfvarson[8,9], Katherine Hanigan[6], Paul Henderson[10,11], Jean-Pierre Hugot[12,13], Amir Karban[14], Nicholas A Kennedy[15], Mohammed Azam Khan[16], Marc Lémann[17], Arie Levine[18], Dunecan Massey[19], Monica Milla[20], Grant W Montgomery[21], Sok Meng Evelyn Ng[22], Ioannis Oikonomou[22], Harald Peeters[23], Deborah D. Proctor[22], Jean-Francois Rahier[24], Rebecca Roberts[2], Paul Rutgeerts[25], Frank Seibold[26], Laura Stronati[27], Kirstin M Taylor[28], Leif Törkvist[29], Kullak Ublick[30], Johan Van Limbergen[31], Andre Van Gossum[32], Morten H. Vatn[33], Hu Zhang[20], Wei Zhang[22], Australia and New Zealand IBDGC*, Belgium Genetic Consortium†, Initiative on Crohn and Colitis, NIDDK IBDGC‡, United Kingdom IBDGC, Wellcome Trust Case Control Consortium§

[1]Department of Medicine, University of Otago, Christchurch, New Zealand. [2]Gastroenterology Unit, INSERM U954, Nancy University and Hospital, France. [3]INSERM, U1019, Lille, France. [4]Univ Lille Nord de France, CHU Lille and Lille-2 University,Gastroenterology Unit, France. [5]Division of Internal Medicine, Villa Sofia-V. Cervello Hospital, University of Palermo, Palermo, Italy. [6]Inflammatory Bowel Diseases, Genetic Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia. [7]Department of Surgical and Gastroenterological Sciences, University of Padua, Padua, Italy. [8]Department of Medicine, Örebro University Hospital, Örebro, Sweden. [9]School of Health and Medical Sciences, Örebro University, Örebro, Sweden. [10]Royal Hospital for Sick Children, Paediatric Gastroenterology and Nutrition, Edinburgh, UK. [11]Child Life and Health, University of Edinburgh, Edinburgh, UK. [12]INSERM U843, Paris, France. [13]Univ-Paris Diderot Sorbonne Paris-Cité, Paris France. [14]Department of Gastroenterology, Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel. [15]Gastrointestinal Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [16]Genetic Medicine, MAHSC, University of Manchester, Manchester, UK. [17]Université Paris Diderot, GETAID group, Paris, France. [18]Pediatric Gastroenterology Unit, Wolfson Medical Center and Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel. [19]Inflammatory Bowel Disease Research Group, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK. [20]Azienda Ospedaliero Universitaria (AOU) Careggi, Unit of Gastroenterology SOD2, Florence, Italy. [21] Molecular Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia. [22]Department of Internal Medicine, Section of Digestive Diseases, Yale School of Medicine, New Haven, Connecticut, USA. [23]Dept Gastroenterology - University hospital Gent - De Pintelaan - 9000 Gent Belgium. [24]Dept Gastroenterology - UCL Mont Godinne Belgium. [25]Division of Gastroenterology, University Hospital Gasthuisberg, Leuven, Belgium. [26]University of Bern, Division of Gastroenterology, Inselspital, Bern,

Switzerland. [27]Department of Radiobiology and Human Health, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Rome, Italy. [28]Dept Gastroenterology, Guy's & St Thomas' NHS Foundation Trust, St Thomas' Hospital, London, UK. [29]Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden. [30]Division of Clinical Pharmacology and Toxicology, University Hospital Zurich, Zurich, Switzerland. [31]Division of Pediatric Gastroenterology, Hepatology and Nutrition, Hospital for Sick Children, Toronto, Ontario, Canada. [32]Dept Gastroenterology - 3University Brussels. [33] Department of Transplantation Medicine, Division of Cancer medicine, Surgery and Transplantation, Oslo University Hospital Rikshospitalet, Oslo, Norway.

*Australia and New Zealand IBDGC

Jane M. Andrews[1], Peter A. Bampton[2], Murray Barclay[3], Timothy H. Florin[4], Richard Gearry[3], Krupa Krishnaprasad[5], Ian C. Lawrance[6], Gillian Mahy[7], Grant W. Montgomery[8], Graham Radford-Smith[5,9], Rebecca L. Roberts[10], Lisa A. Simms[5].

[1]Inflammatory Bowel Disease Service, Department of Gastroenterology and Hepatology, Royal Adelaide Hospital, and School of Medicine, University of Adelaide, Adelaide, Australia. [2]Department of Gastroenterology and Hepatology, Flinders Medical Centre and School of Medicine, Flinders University, Adelaide, Australia. [3]Department of Gastroenterology, Christchurch Hospital and Department of Medicine, University of Otago, Christchurch, New Zealand. [4]Department of Gastroenterology, Mater Health Services, Brisbane, Australia, and School of Medicine, University of Queensland, Brisbane, Australia. [5]Inflammatory Bowel Diseases, Genetics and Computational Biology, Queensland Institute of Medical Research, Brisbane, Australia. [6]Centre for Inflammatory Bowel Diseases, Fremantle Hospital and School of Medicine and Pharmacology, The University of Western Australia, Fremantle, Australia. [7]Department of Gastroenterology, The Townsville Hospital and James Cook University School of Medicine, Townsville, Australia. [8]Molecular Epidemiology, Genetics and Computational Biology, Queensland Institute of Medical Research, Brisbane, Australia. [9]Department of Gastroenterology, Royal Brisbane and Womens Hospital, and School of Medicine, University of Queensland, Brisbane, Australia. [10]University of Otago, Department of Medicine, Christchurch, New Zealand.

†Belgium Genetic Consortium

Leila Amininijad[1], Isabelle Cleynen[2], Olivier Dewit[3], Denis Franchimont[1], Michel Georges[4], Debby Laukens[5], Harald Peeters[5], Jean-Francois Rahier[3], Paul Rutgeerts[2], Emilie Theatre[4], [6], André Van Gossum[1], Severine Vermeire[7].

[1]Erasmus Hospital, Free University of Brussels, Department of Gastroenterology, Brussels, Belgium. [2]Department of Pathophysiology, Gastroenterology section, KU Leuven, Leuven, Belgium. [3]Department of Gastroenterology, Clinique Universitaire St-

Luc, Brussels, Belgium. [4]Unit of Animal Genomics, Groupe Interdisciplinaire de Gnoprotomique Applique (GIGA-R) and Faculty of Veterinary Medicine, University of Lige, Lige, Belgium. [5]Ghent University Hospital, Department of Gastroenterology and Hepatology, Ghent, Belgium. [6]Division of Gastroenterology, Centre Hospitalier Universitaire, Universit de Lige, Lige, Belgium. [7]Division of Gastroenterology, University Hospital Gasthuisberg, Leuven, Belgium.

‡NIDDK Inflammatory Bowel Disease Genetics Consortium‡

Guy Aumais[1], Leonard Baidoo[2], Arthur M. Barrie III[2], Karen Beck[2], Edmond-Jean Bernard[3], David G. Binion[2], Alain Bitton[4], Steve R. Brant[5], Judy H. Cho[6,7], Albert Cohen[8], Kenneth Croitoru[9], Mark J. Daly[10,11], Lisa W. Datta[5], Colette Deslandres[12], Richard H. Duerr[2,13], Debra Dutridge[14], John Ferguson[7], Joann Fultz[2], Philippe Goyette[15], Gordon R. Greenberg[9], Talin Haritunians[14], Gilles Jobin[16], Seymour Katz[17], Raymond G. Lahaie[18], Dermot P. McGovern[14,19], Linda Nelson[2], Sok Meng Ng[7], Kaida Ning[7], Ioannis Oikonomou[7], Pierre Paré[20], Deborah D. Proctor[7], Miguel D. Regueiro[2], John D. Rioux[15], Elizabeth Ruggiero[7], L. Philip Schumm[21], Marc Schwartz[2], Regan Scott[2], Yashoda Sharma[7], Mark S. Silverberg[9], Denise Spears[5], A. Hillary Steinhart[9], Joanne M. Stempak[9], Jason M. Swoger[2], Constantina Tsagarelis[4], Wei Zhang[7], Clarence Zhang[22], Hongyu Zhao[22].

[1]University of Montreal, Maissonneuve – Rosemont Hospital, Quebec Association of Gastroenterologists, Montréal, Québec, Canada. [2]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [3]Hôpital Hôtel Dieu, Montréal, Québec, Canada. [4]Division of Gastroenterology, McGill University Health Centre, Royal Victoria Hospital, Montréal, Québec, Canada. [5]Inflammatory Bowel Disease Center, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [6]Department of Genetics, Yale School of Medicine, New Haven, Connecticut, USA. [7]Department of Internal Medicine, Section of Digestive Diseases, Yale School of Medicine, New Haven, Connecticut, USA. [8]Division of Gastroenterology, Hôpital Général Juif Sir Mortimer B. Davis Jewish General Hospital, Montréal, Québec, Canada. [9]Mount Sinai Hospital Inflammatory Bowel Disease Centre, University of Toronto, Toronto, Ontario, Canada. [10]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA. [11]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [12]Hopital Sainte Justine, Montréal, Québec, Canada. [13]Department of Human Genetics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania, USA. [14]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [15]Université de Montréal and the Montreal Heart Institute, Research Center, Montréal, Québec, Canada. [16]Pavillon Maisonneuve, Montréal, Québec, Canada. [17]Long Island Clinical Research Associates, Great Neck, New York, USA. [18]CHUM – Hopital Sainte-Luc, Montréal,

Québec, Canada. [19]Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [20]Laval University, Quebec City, Québec, Canada. [21]Department of Health Studies, University of Chicago, Chicago, Illinois, USA. [22]Department of Biostatistics, School of Public Health, Yale University, New Haven, Connecticut, USA.

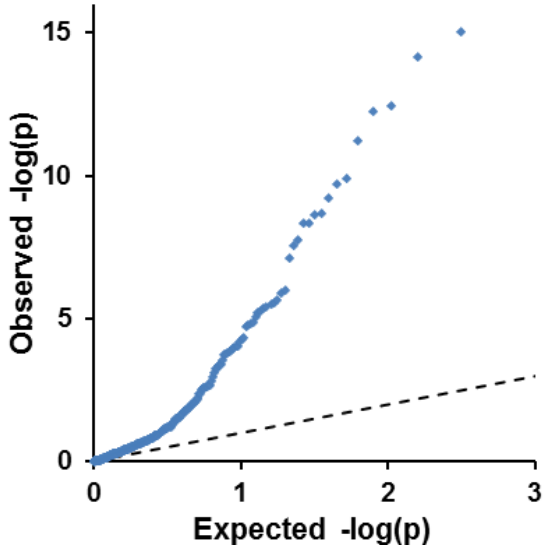
Corresponding authors

WTCCC - Peter Donnelly, Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG, UK.

IIBDGC - Jeff C. Barrett, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA UK.

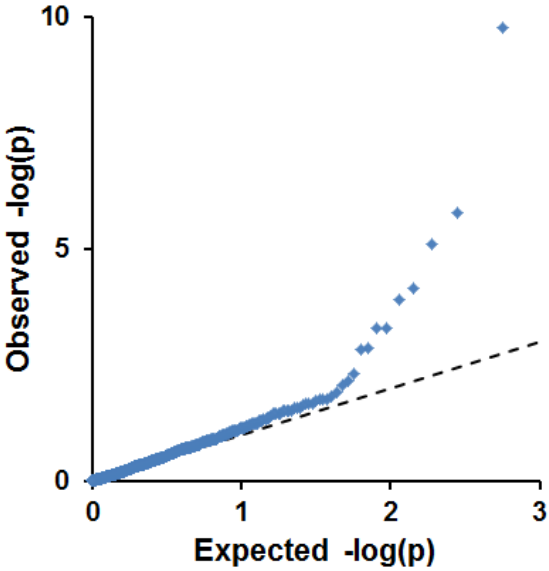
Supplementary Figures

Supplementary Figure 1. QQ plot of the ImmunoChip results nominated from our 2011 GWAS for deep replication.



This consisted of 317 independent SNPs mapping outside of fine mapping intervals.[10]

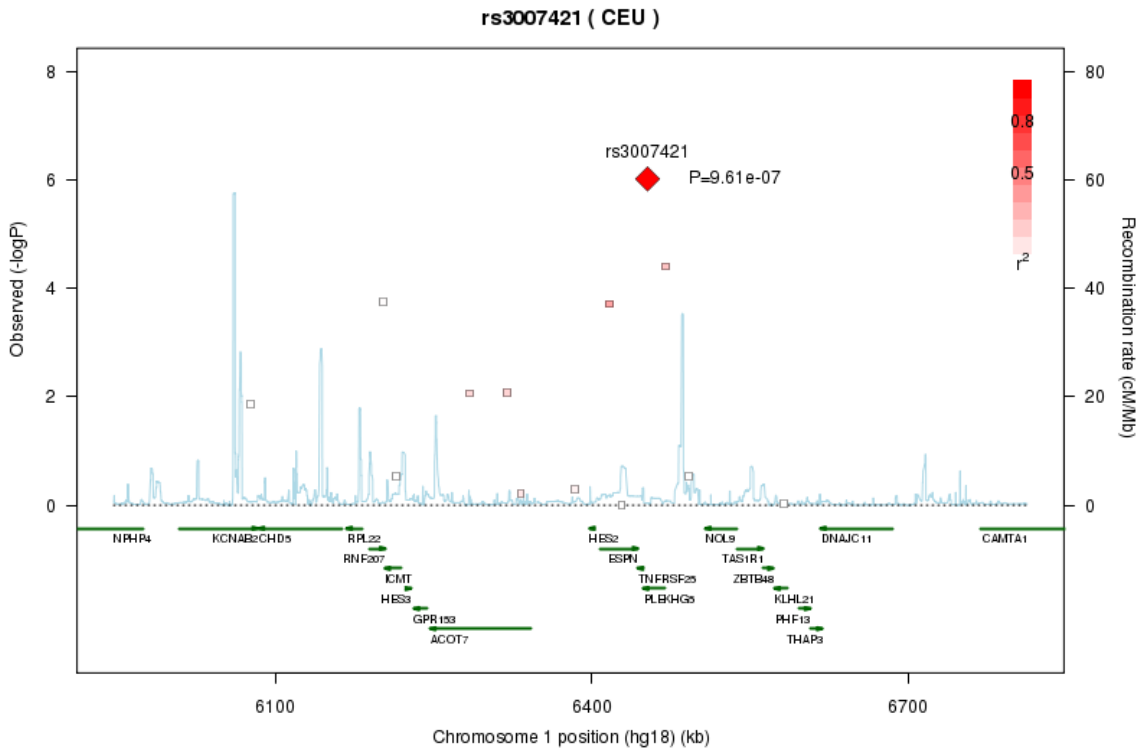
Supplementary Figure 2. QQ plot for the 567 independent SNPs nominated from our meta-analysis of GWAS.



The meta-analysis of GWAS[11] was published prior to our main GWAS.[10] Note the peak SNP on this figure is rs9282641 (CD86), this is the only SNP shared with the SNPs nominated from the main GWAS. Association with this SNP was reported in the main GWAS.[10]

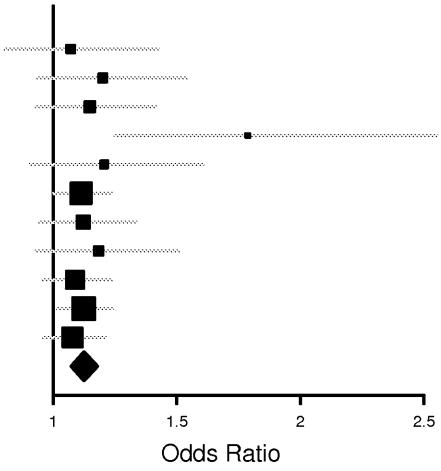
Supplementary Figure 3. Discovery phase rs3007421.

A



B

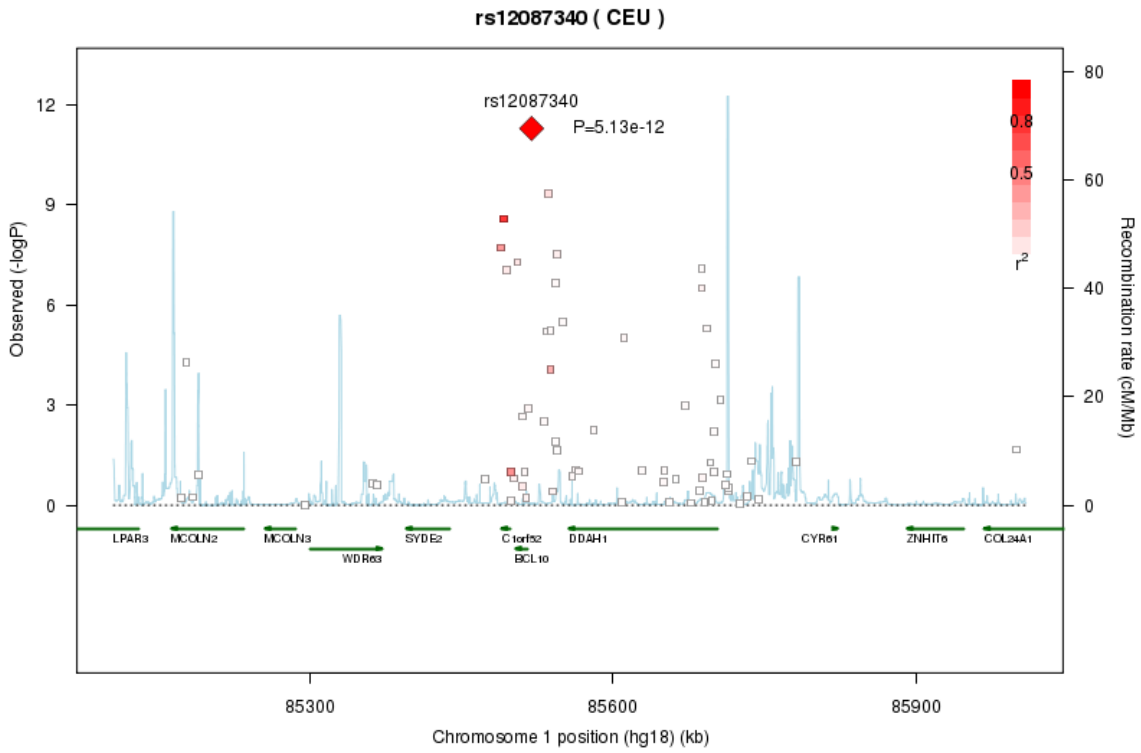
Stratum	OR	95% CI
AUSNZ	1.07	[0.80–1.43]
Belgium	1.20	[0.94–1.54]
Denmark	1.15	[0.93–1.41]
Finland	1.79	[1.25–2.55]
France	1.21	[0.91–1.60]
Germany	1.11	[1.00–1.24]
Italy	1.12	[0.94–1.33]
Norway	1.18	[0.93–1.51]
Sweden	1.09	[0.96–1.24]
UK	1.12	[1.02–1.24]
US	1.08	[0.96–1.21]
Summary	**1.12**	**[1.07–1.18]**



A) Regional Association and B) Forest Plot

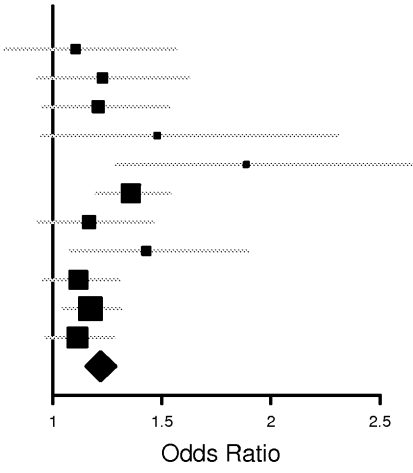
Supplementary Figure 4. Discovery phase rs12087340.

A



B

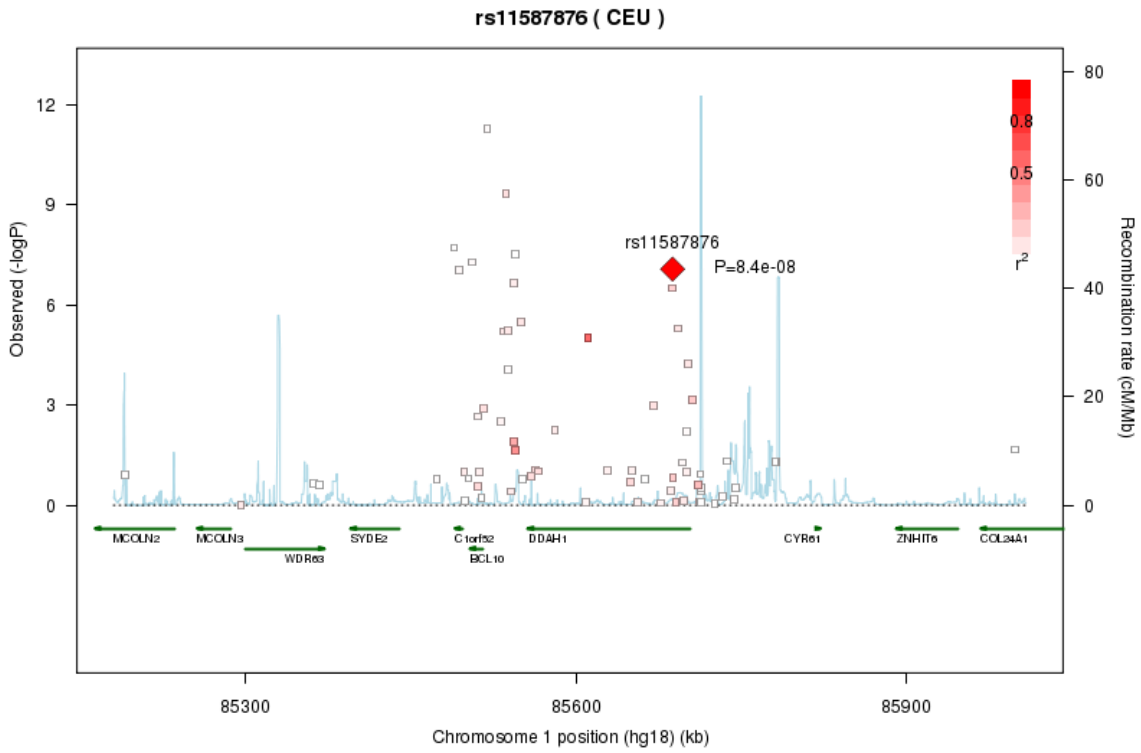
Stratum	OR	95% CI
AUSNZ	1.10	[0.78–1.56]
Belgium	1.23	[0.93–1.62]
Denmark	1.21	[0.95–1.53]
Finland	1.48	[0.95–2.31]
France	1.89	[1.29–2.76]
Germany	1.36	[1.20–1.54]
Italy	1.17	[0.93–1.46]
Norway	1.43	[1.08–1.89]
Sweden	1.12	[0.96–1.30]
UK	1.17	[1.04–1.31]
US	1.11	[0.97–1.28]
Summary	**1.22**	**[1.15–1.29]**



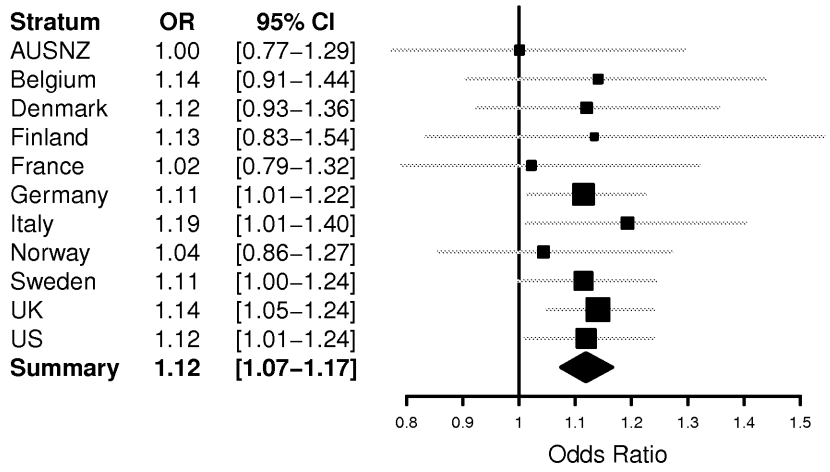
A) Regional Association and B) Forest Plot

Supplementary Figure 5. Discovery phase rs11587876.

A



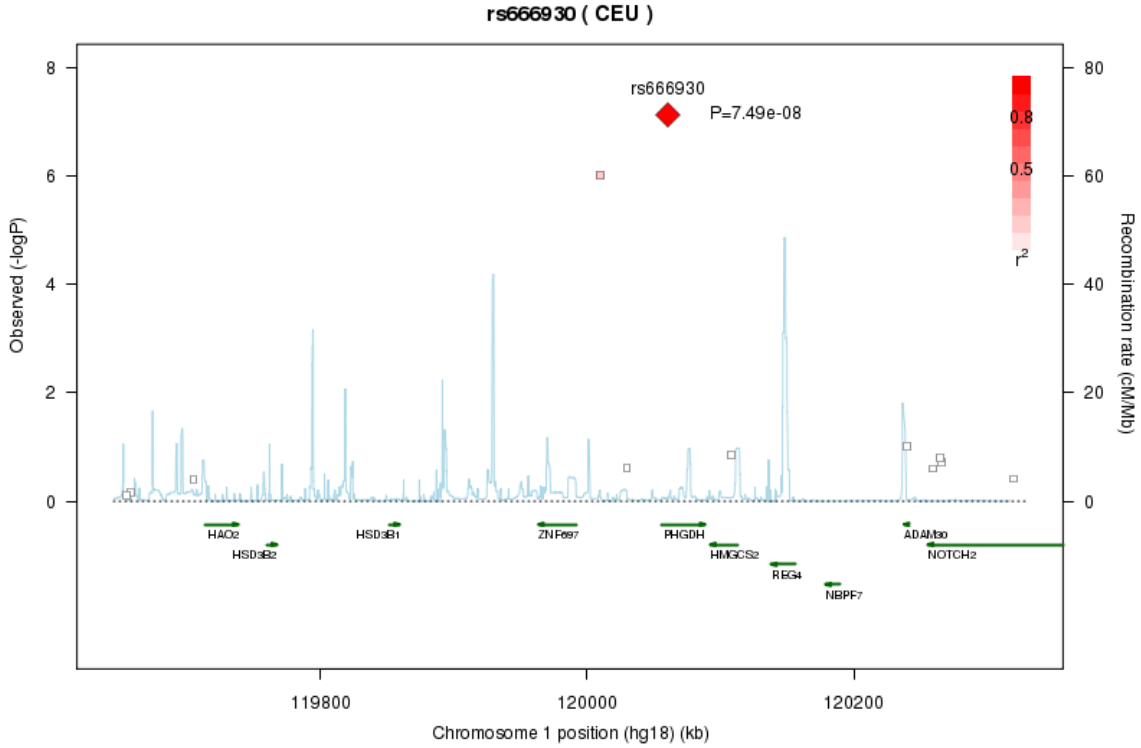
B



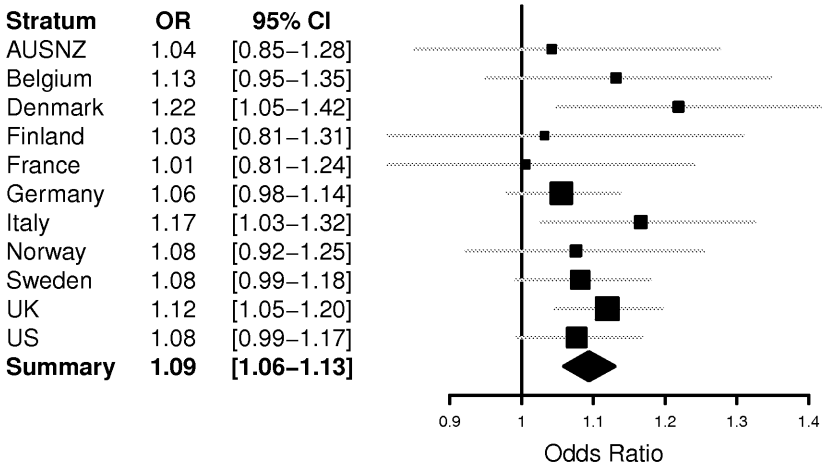
A) Regional Association and B) Forest Plot

Supplementary Figure 6. Discovery phase rs666930.

A



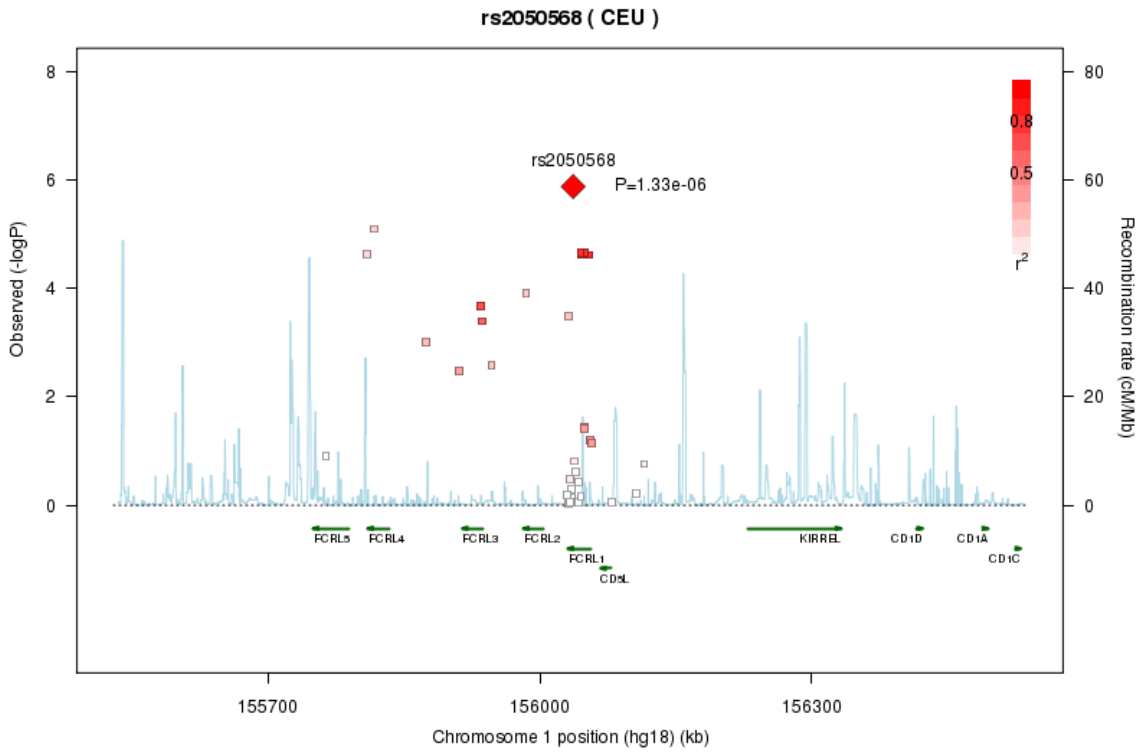
B



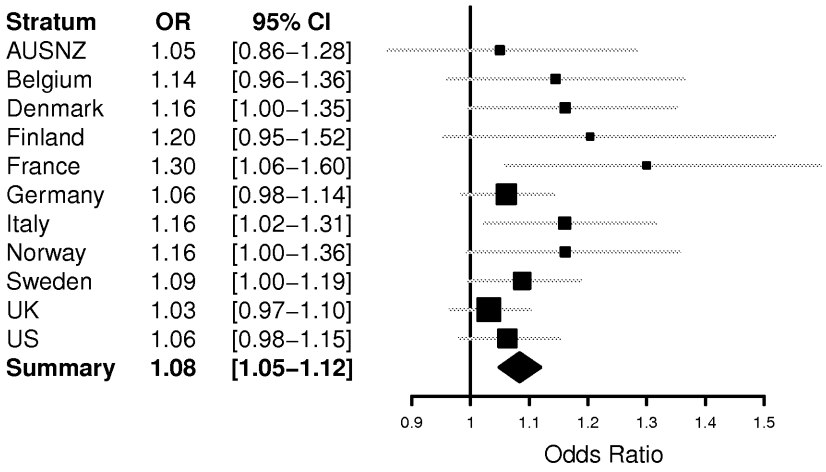
a) Regional Association and b) Forest Plot

Supplementary Figure 7. Discovery phase rs2050568.

A



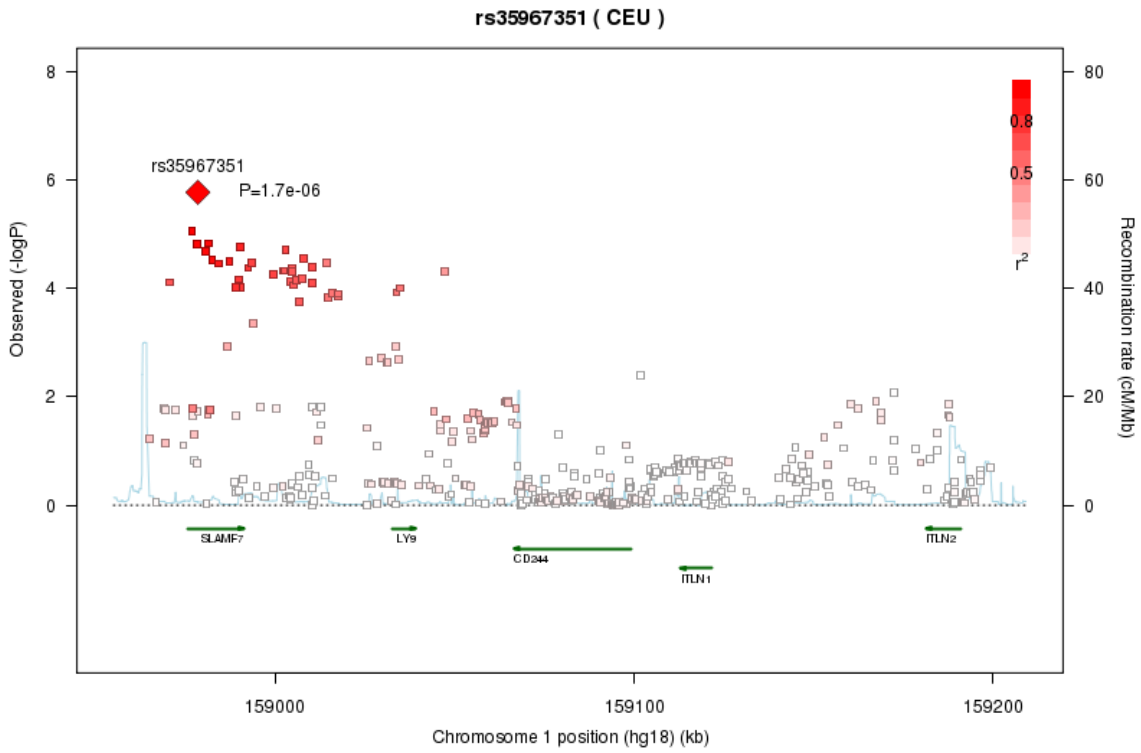
B



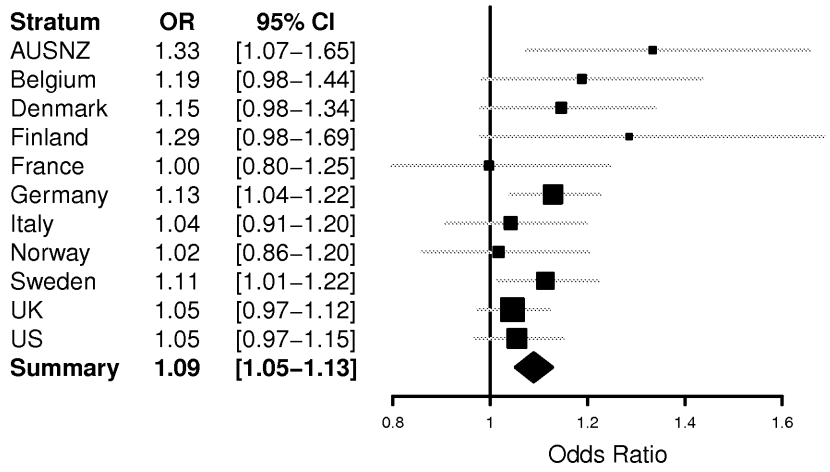
A) Regional Association and B) Forest Plot

Supplementary Figure 8. Discovery phase rs35967351.

A



B

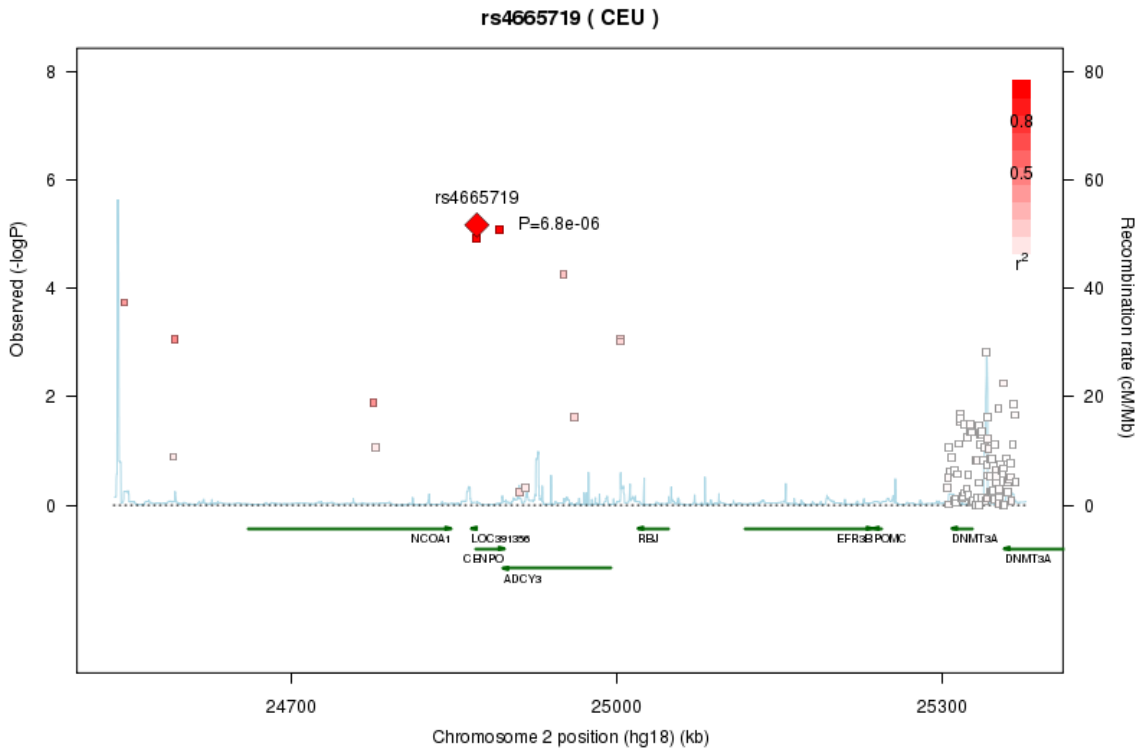


Stratum	OR	95% CI
AUSNZ	1.33	[1.07–1.65]
Belgium	1.19	[0.98–1.44]
Denmark	1.15	[0.98–1.34]
Finland	1.29	[0.98–1.69]
France	1.00	[0.80–1.25]
Germany	1.13	[1.04–1.22]
Italy	1.04	[0.91–1.20]
Norway	1.02	[0.86–1.20]
Sweden	1.11	[1.01–1.22]
UK	1.05	[0.97–1.12]
US	1.05	[0.97–1.15]
Summary	**1.09**	**[1.05–1.13]**

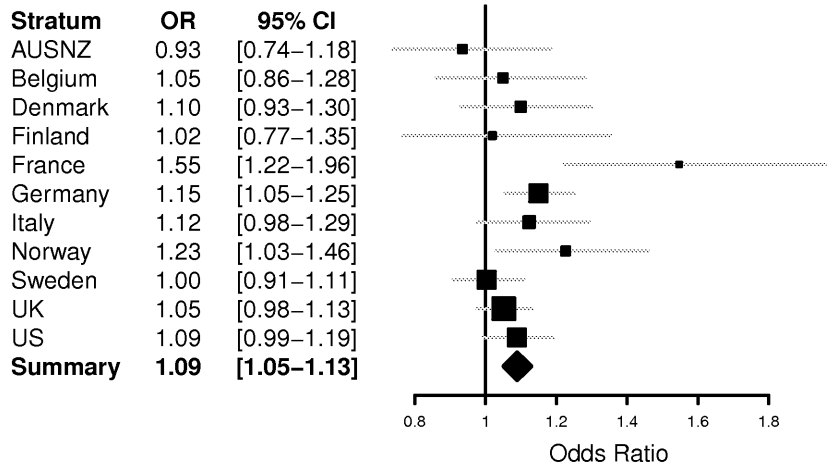
A) Regional Association and B) Forest Plot

Supplementary Figure 9. Discovery phase rs4665719.

A



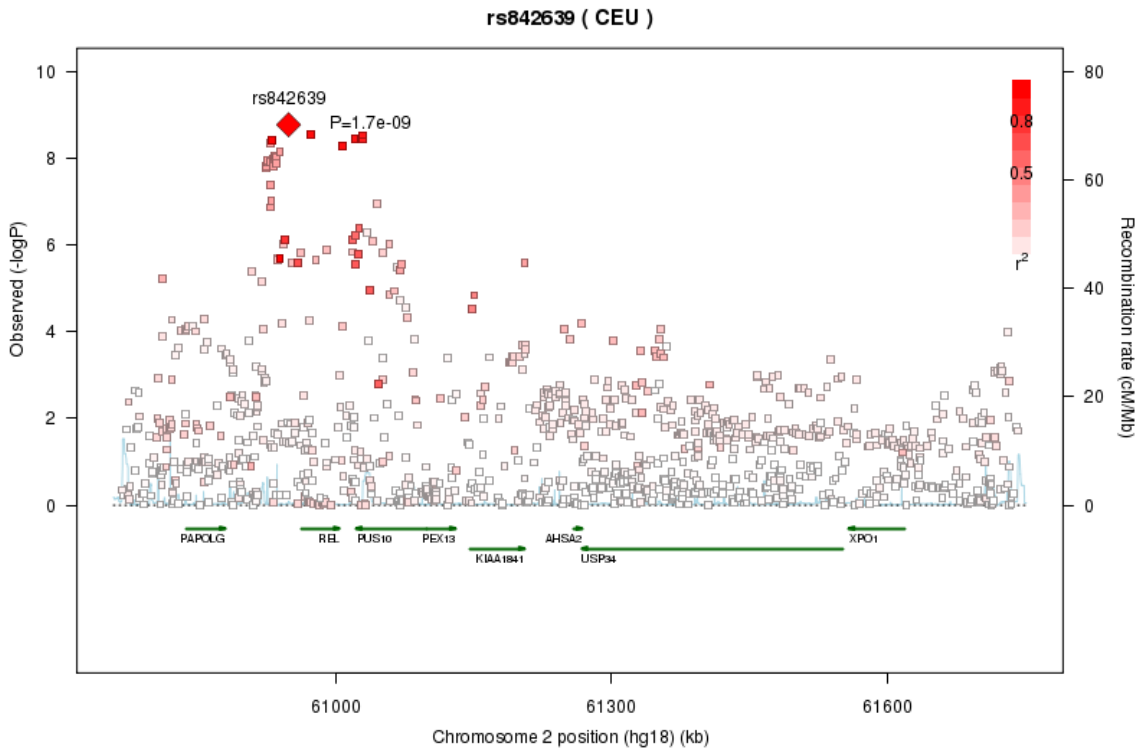
B



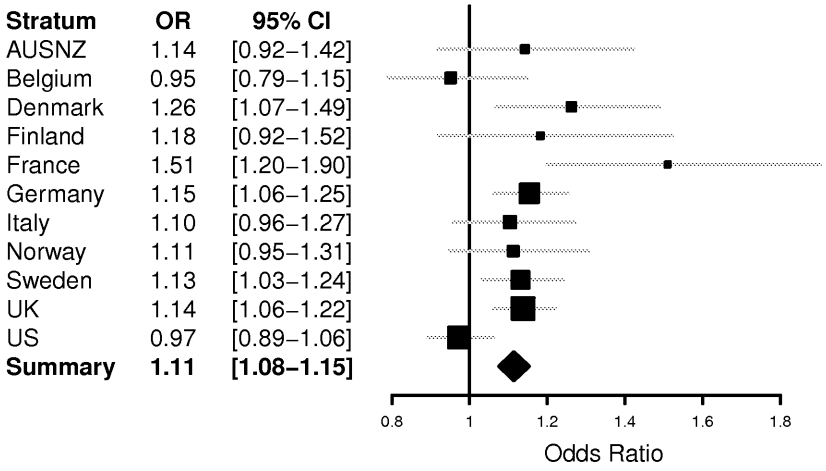
A) Regional Association and B) Forest Plot

Supplementary Figure 10. Discovery phase rs842639.

A



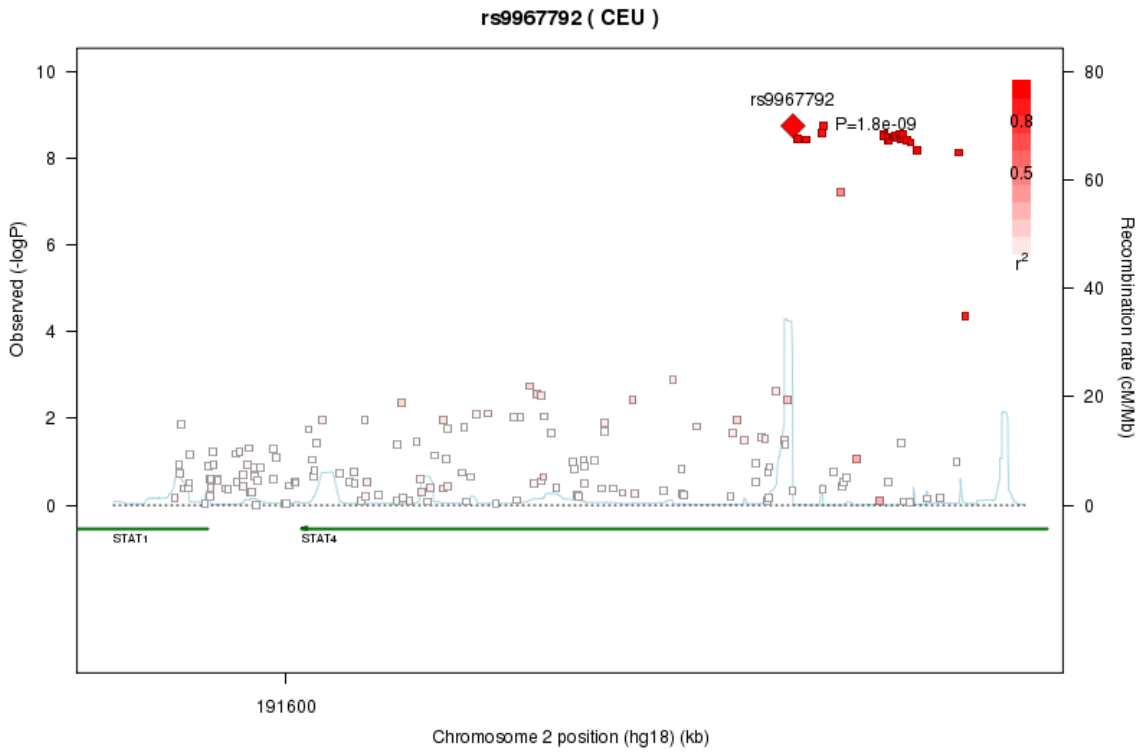
B



A) Regional Association and B) Forest Plot

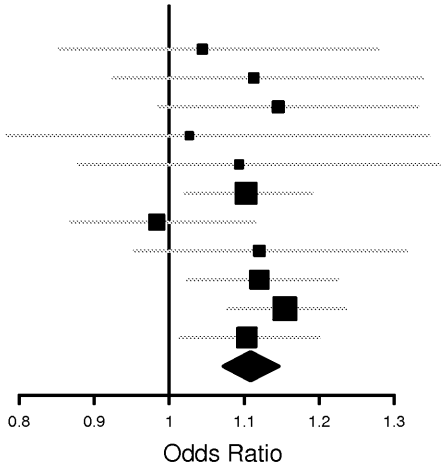
Supplementary Figure 11. Discovery phase rs9967792.

A



B

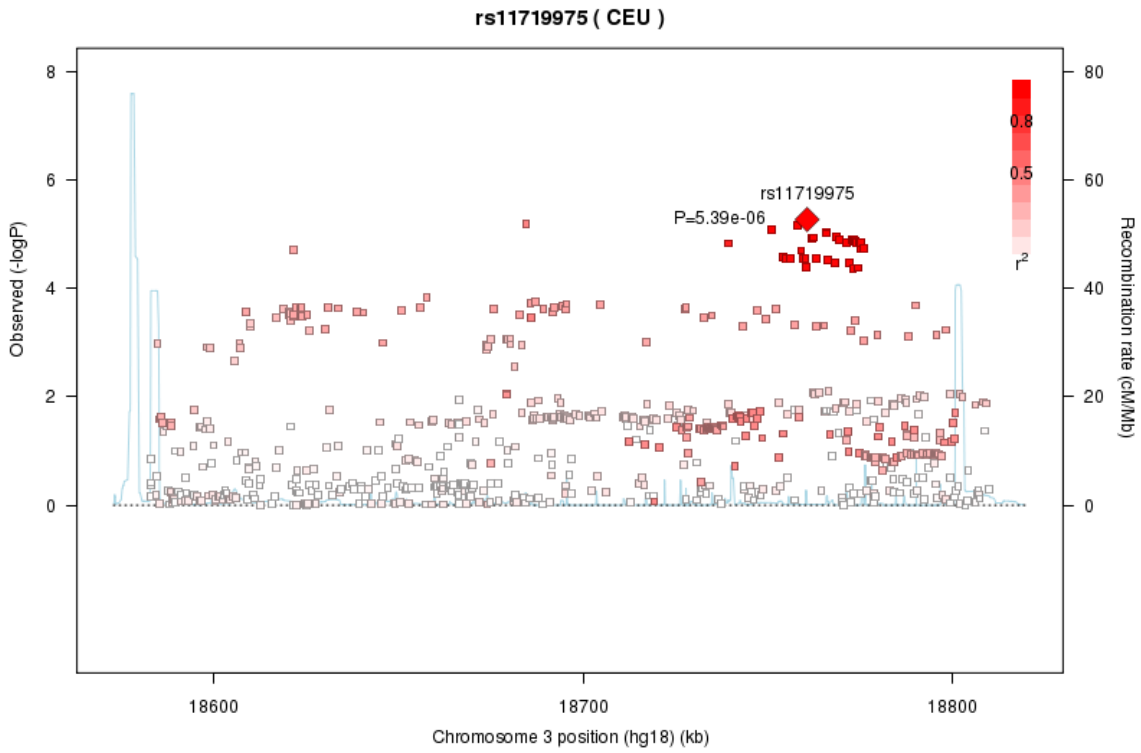
Stratum	OR	95% CI
AUSNZ	1.04	[0.85–1.28]
Belgium	1.11	[0.93–1.34]
Denmark	1.15	[0.99–1.33]
Finland	1.03	[0.78–1.35]
France	1.09	[0.88–1.36]
Germany	1.10	[1.02–1.19]
Italy	0.98	[0.87–1.11]
Norway	1.12	[0.95–1.32]
Sweden	1.12	[1.02–1.22]
UK	1.15	[1.08–1.24]
US	1.10	[1.02–1.20]
Summary	**1.11**	**[1.07–1.15]**



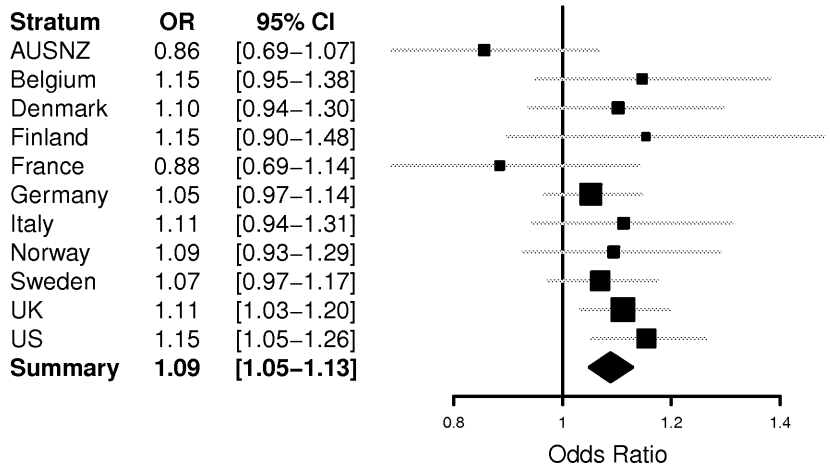
A) Regional Association and B) Forest Plot

Supplementary Figure 12. Discovery phase rs11719975.

A



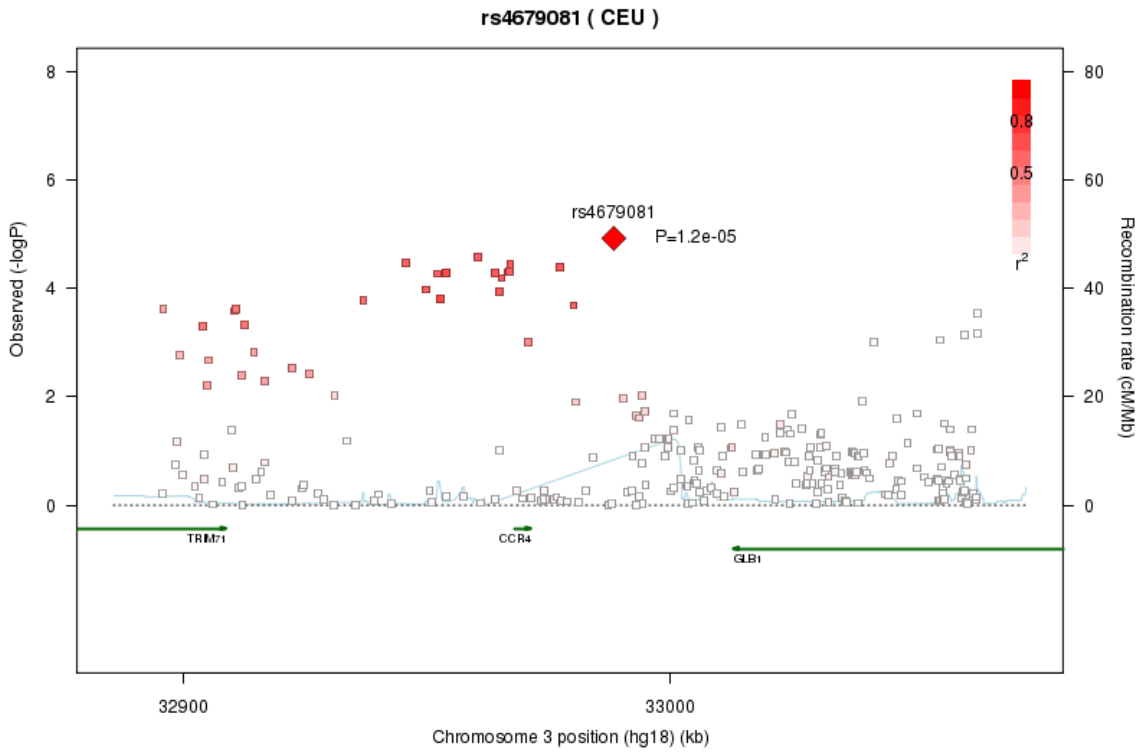
B



A) Regional Association and b) Forest Plot

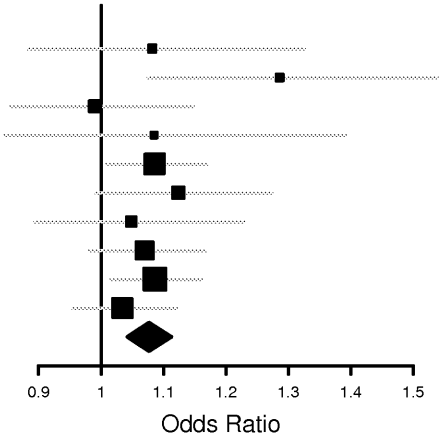
Supplementary Figure 13. Discovery phase rs4679081.

A



B

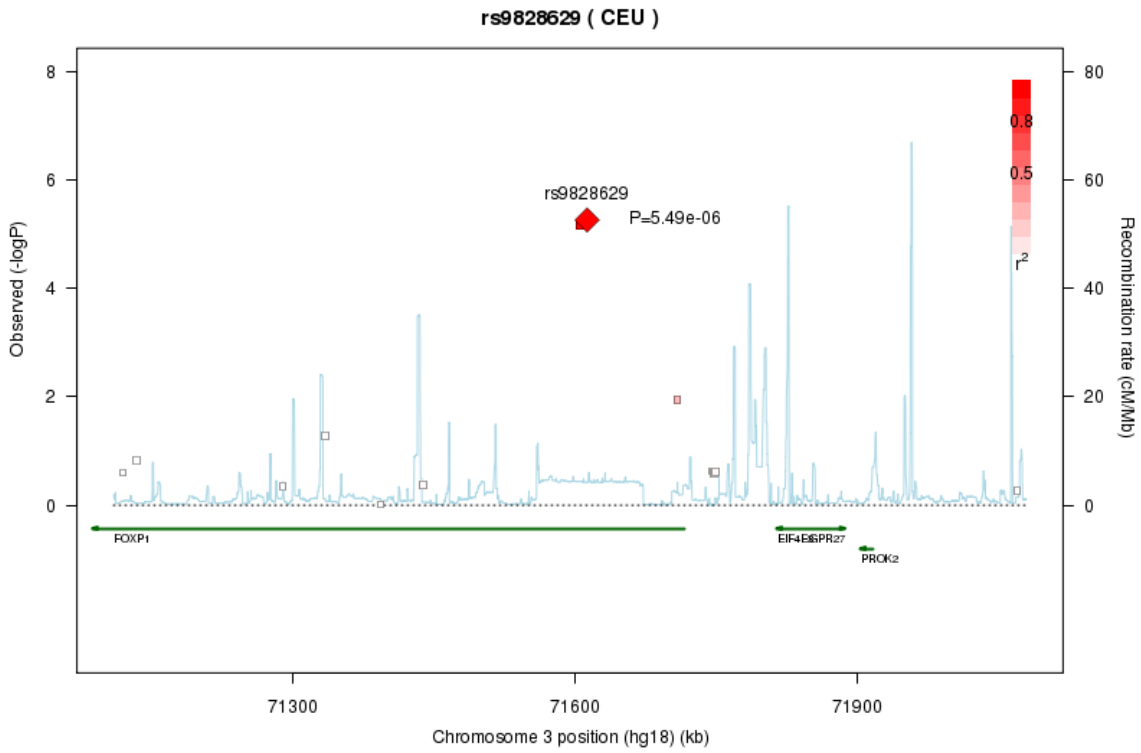
Stratum	OR	95% CI
AUSNZ	1.08	[0.88–1.33]
Belgium	1.29	[1.08–1.54]
Denmark	0.99	[0.85–1.15]
Finland	1.09	[0.85–1.39]
Germany	1.09	[1.01–1.17]
Italy	1.12	[0.99–1.27]
Norway	1.05	[0.89–1.23]
Sweden	1.07	[0.98–1.17]
UK	1.09	[1.02–1.16]
US	1.03	[0.95–1.12]
Summary	**1.08**	**[1.04–1.11]**



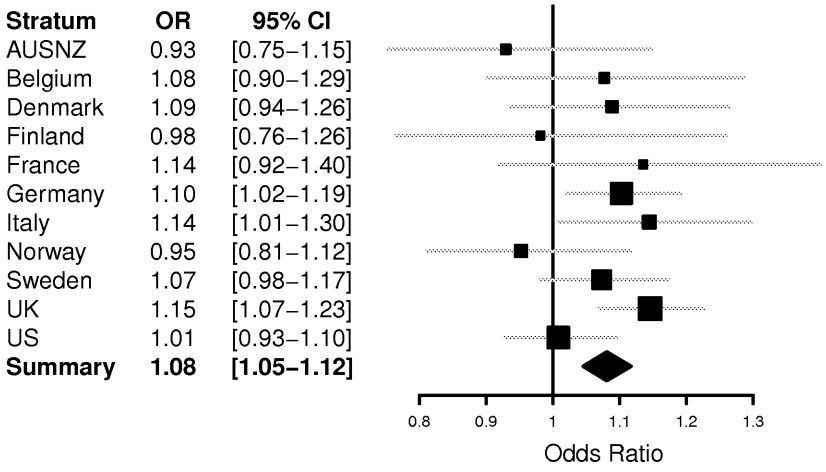
A) Regional Association and b) Forest Plot

Supplementary Figure 14. Discovery phase rs9828629.

A



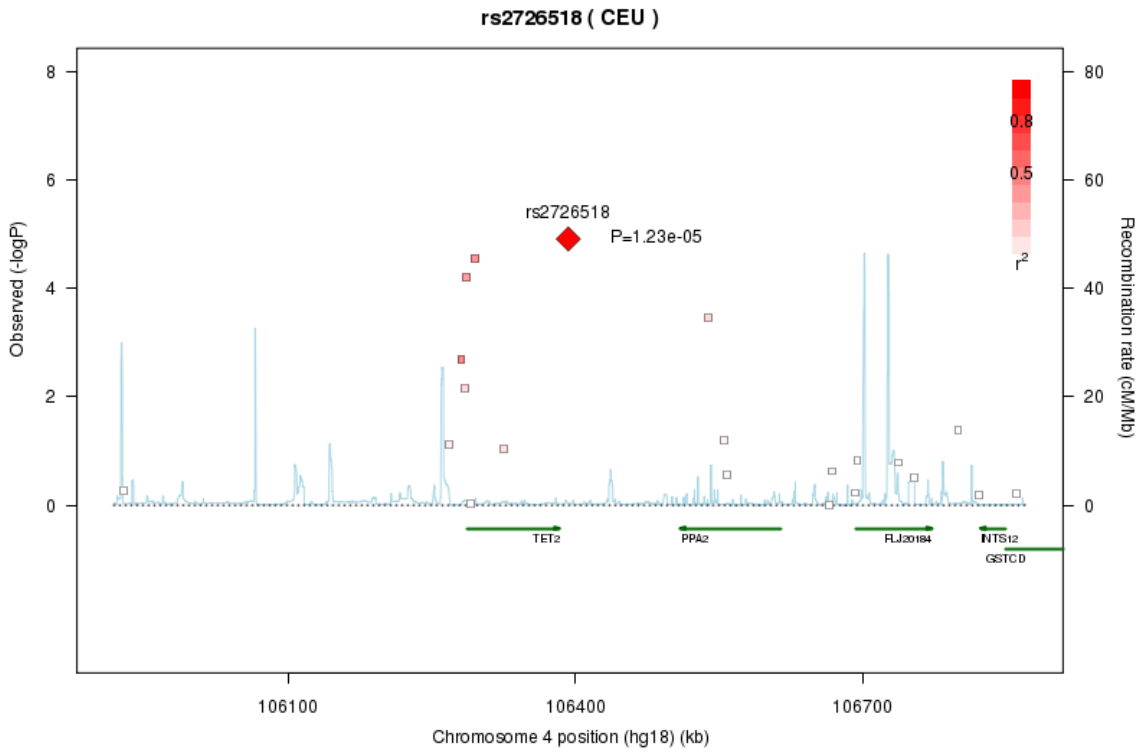
B



A) Regional Association and B) Forest Plot

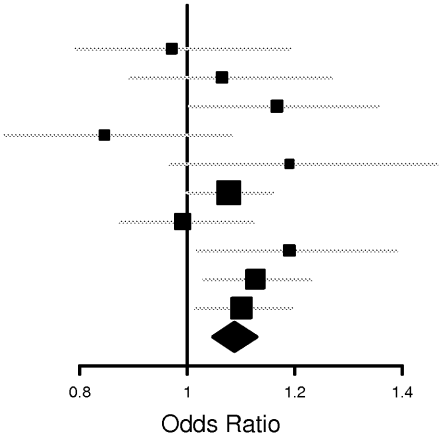
Supplementary Figure 15. Discovery phase rs2726518.

A



B

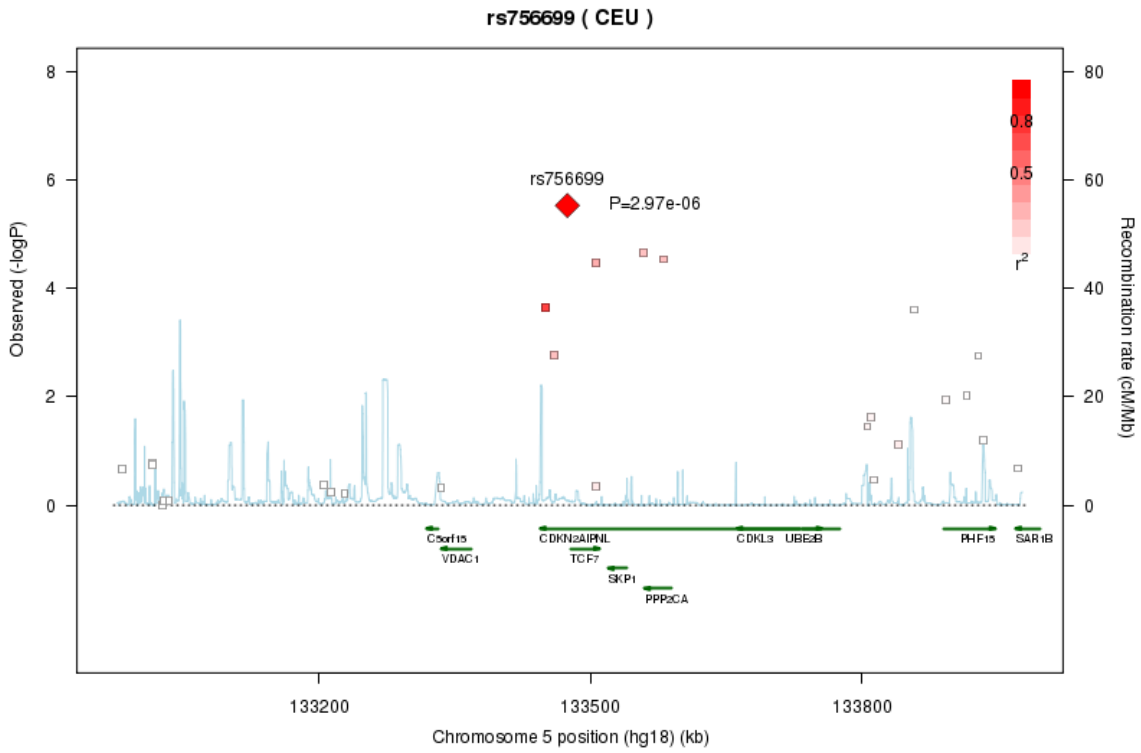
Stratum	OR	95% CI
AUSNZ	0.97	[0.79–1.19]
Belgium	1.06	[0.89–1.27]
Denmark	1.17	[1.00–1.36]
Finland	0.85	[0.66–1.08]
France	1.19	[0.97–1.46]
Germany	1.08	[1.00–1.16]
Italy	0.99	[0.87–1.12]
Norway	1.19	[1.02–1.39]
Sweden	1.13	[1.03–1.23]
US	1.10	[1.01–1.19]
Summary	**1.09**	**[1.05–1.13]**



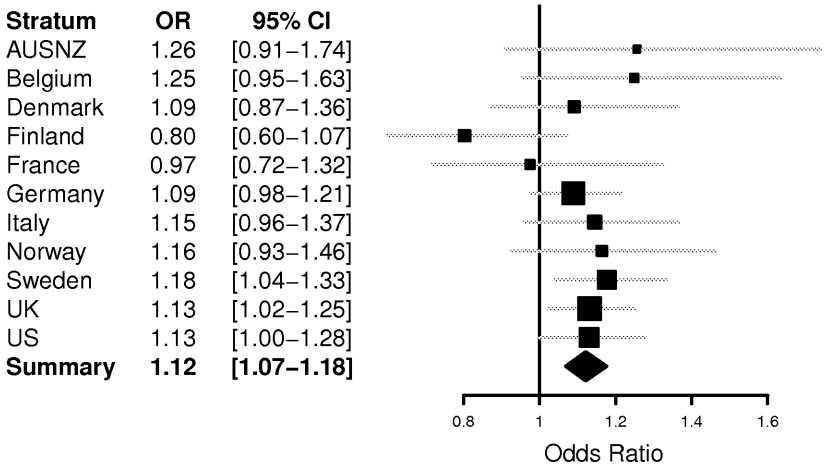
A) Regional Association and b) Forest Plot

Supplementary Figure 16. Discovery phase rs756699.

A



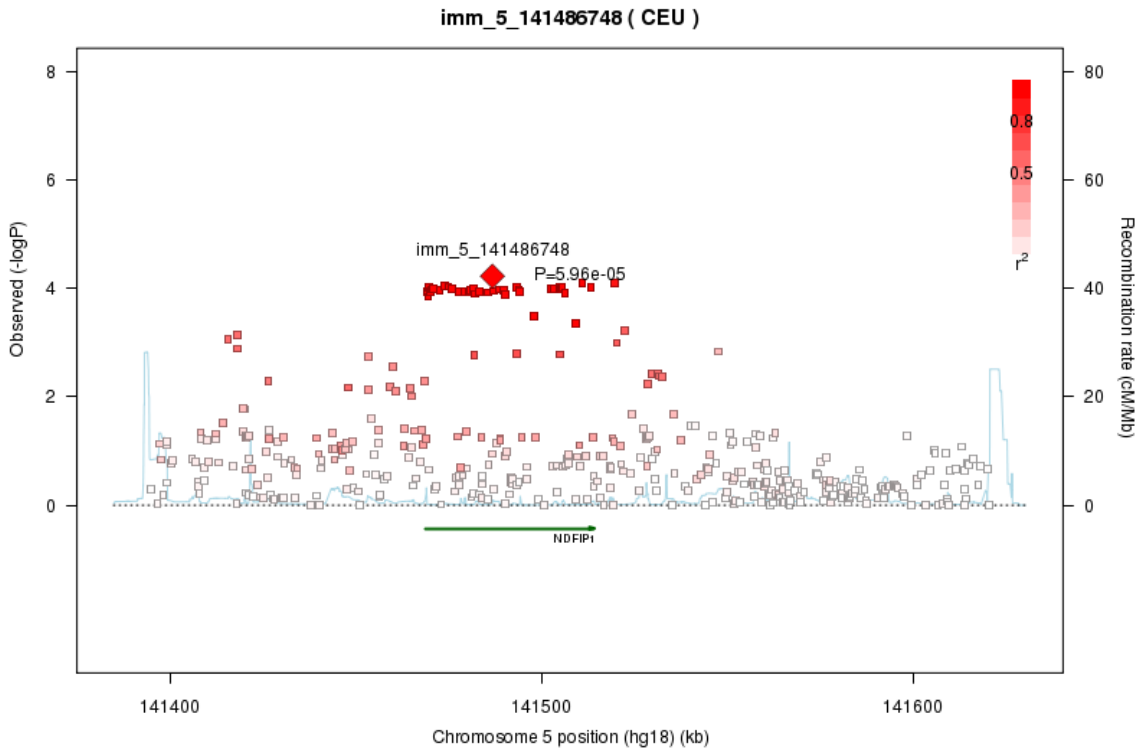
B



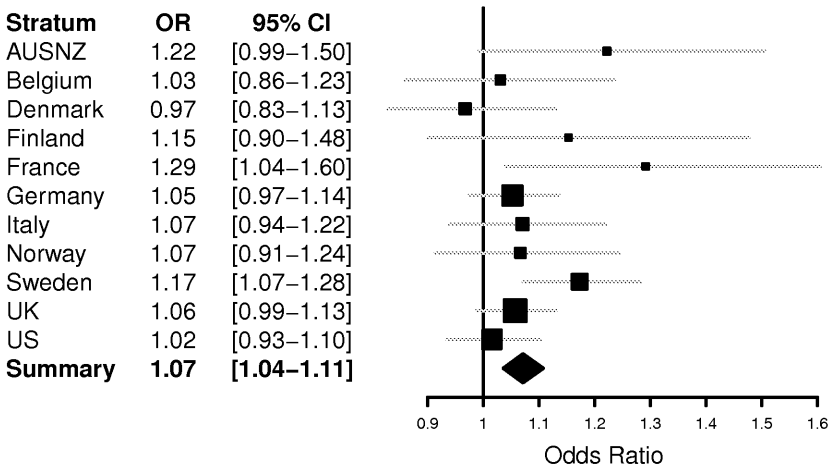
A) Regional Association and B) Forest Plot

Supplementary Figure 17. Discovery phase 5:141506564.

A



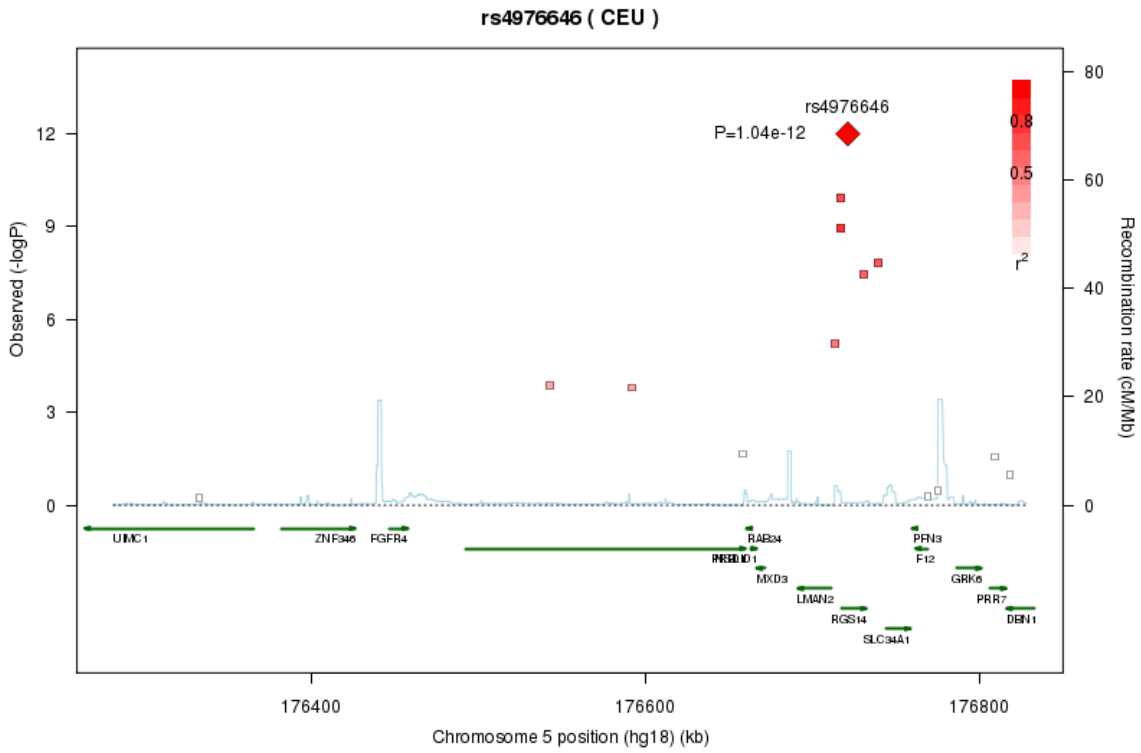
B



A) Regional Association and b) Forest Plot for 5:141506564 (imm_5_141486748)

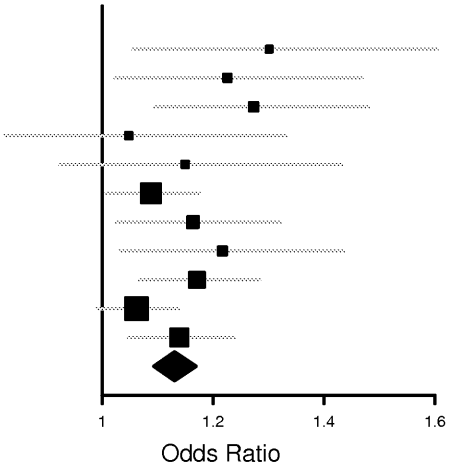
Supplementary Figure 18. Discovery phase rs4976646.

A



B

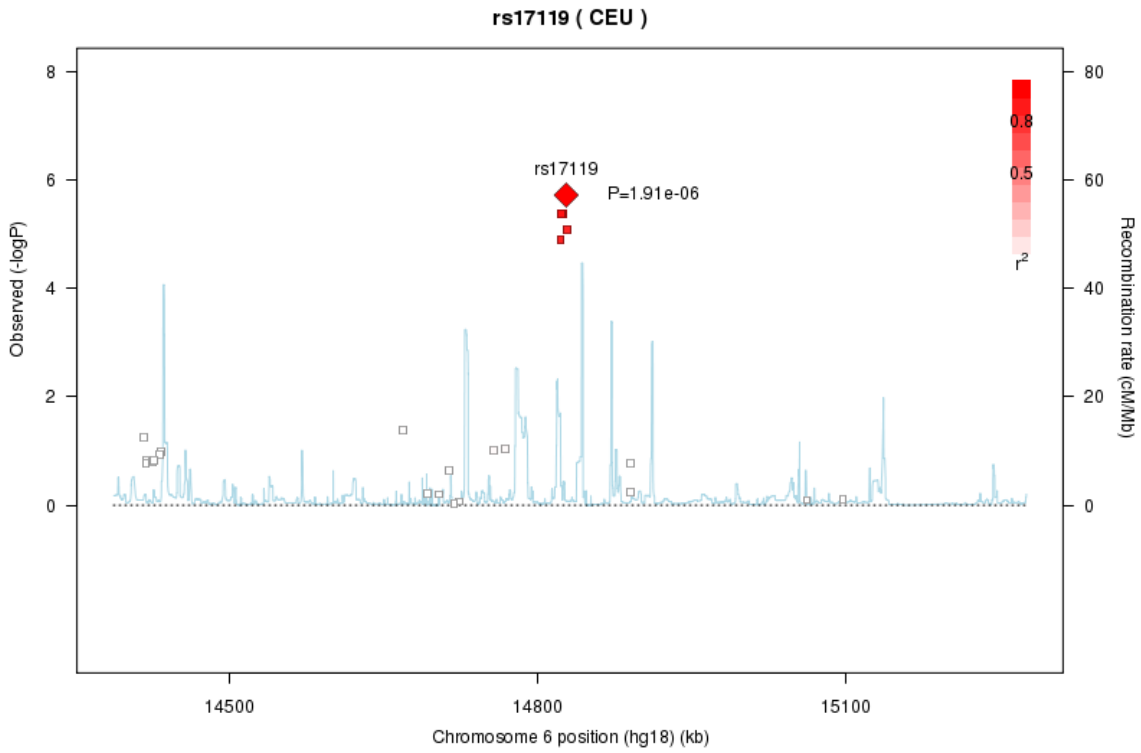
Stratum	OR	95% CI
AUSNZ	1.30	[1.06–1.60]
Belgium	1.23	[1.02–1.47]
Denmark	1.27	[1.09–1.48]
Finland	1.05	[0.82–1.33]
France	1.15	[0.92–1.43]
Germany	1.09	[1.01–1.18]
Italy	1.16	[1.02–1.32]
Norway	1.22	[1.03–1.43]
Sweden	1.17	[1.07–1.28]
UK	1.06	[0.99–1.14]
US	1.14	[1.05–1.24]
Summary	**1.13**	**[1.09–1.17]**



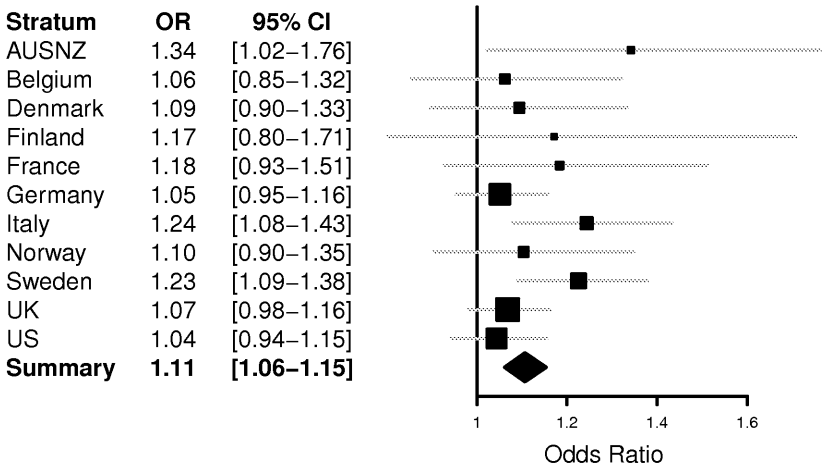
A) Regional Association and B) Forest Plot

Supplementary Figure 19. Discovery phase rs17119.

A



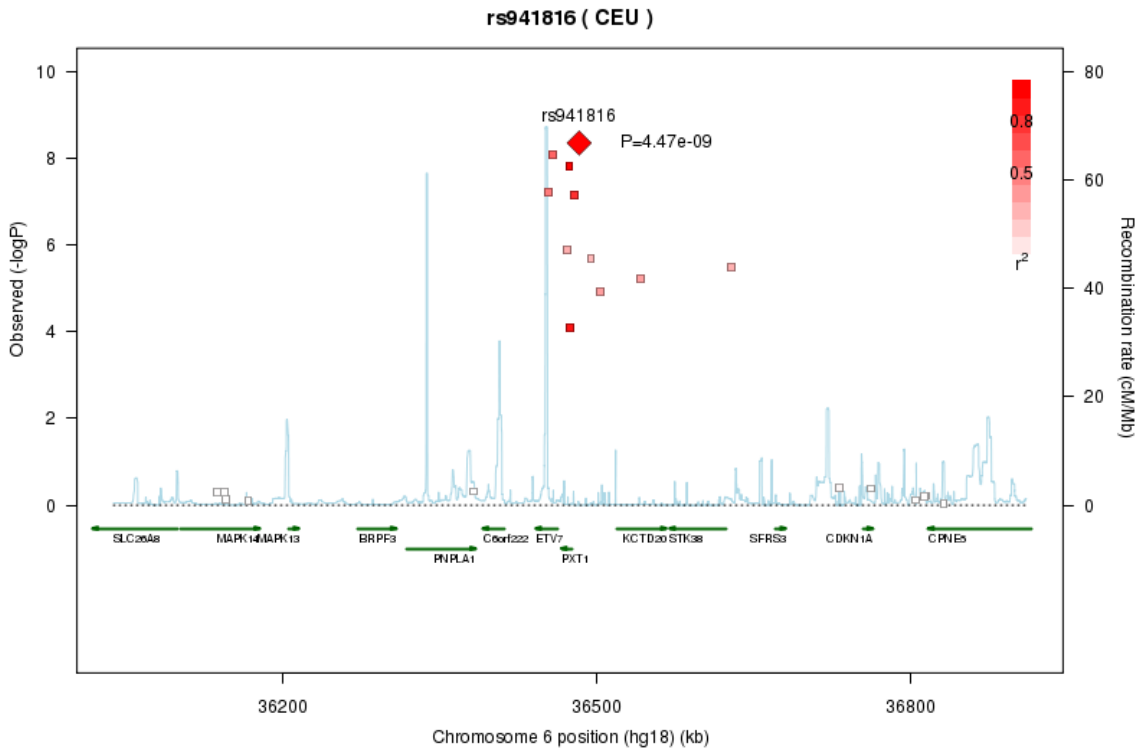
B



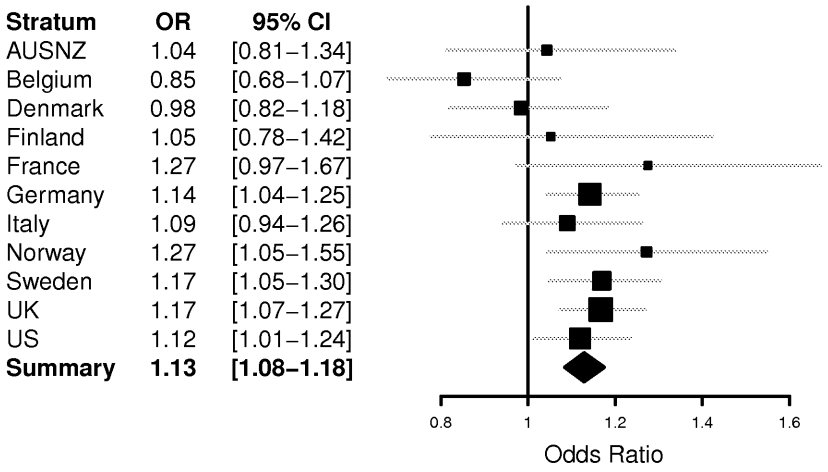
A) Regional Association and B) Forest Plot

Supplementary Figure 20. Discovery phase rs941816.

A



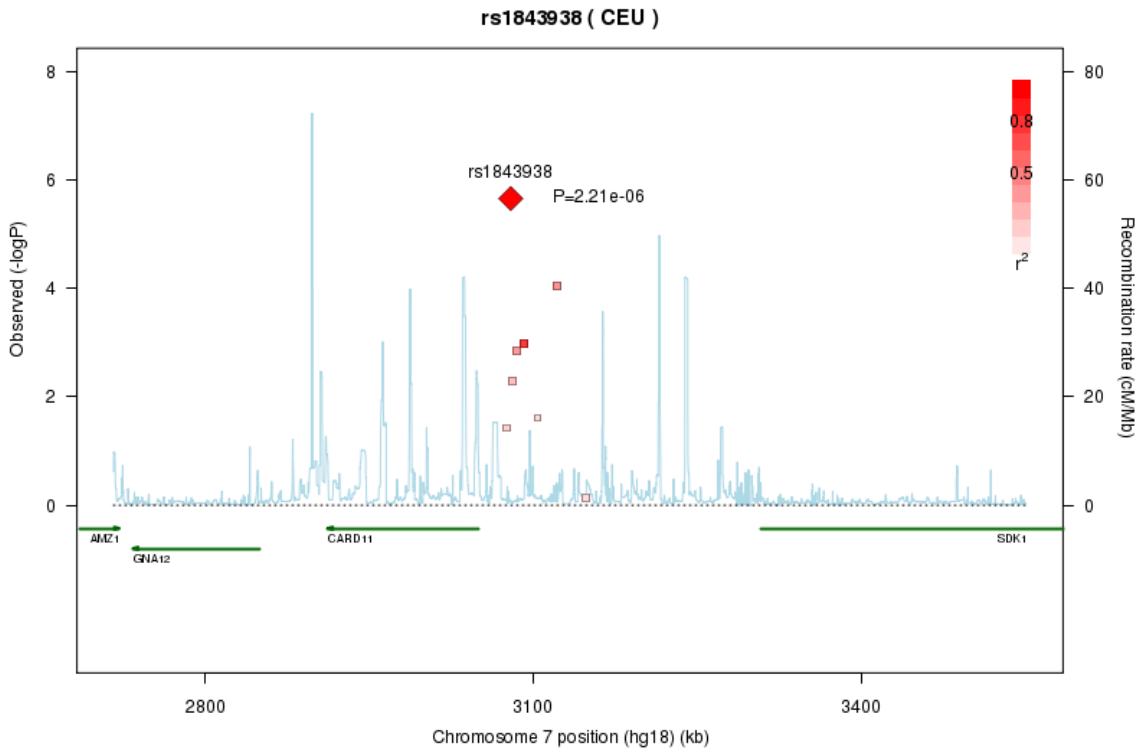
B



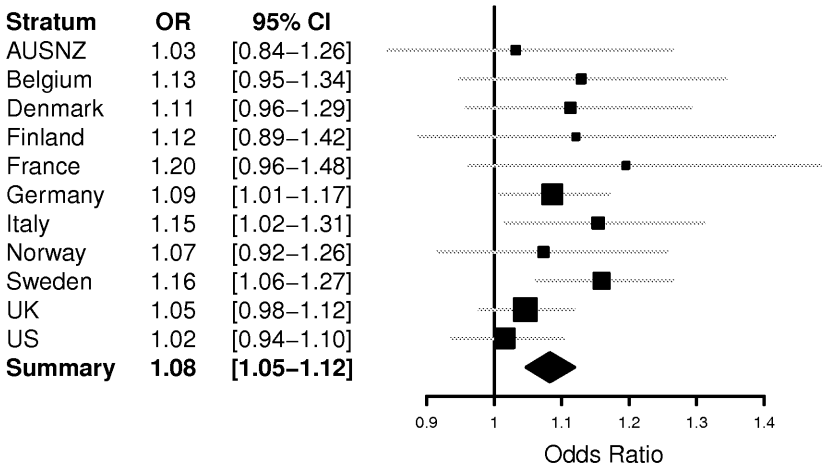
A) Regional Association and B) Forest Plot

Supplementary Figure 21. Discovery phase rs1843938.

A



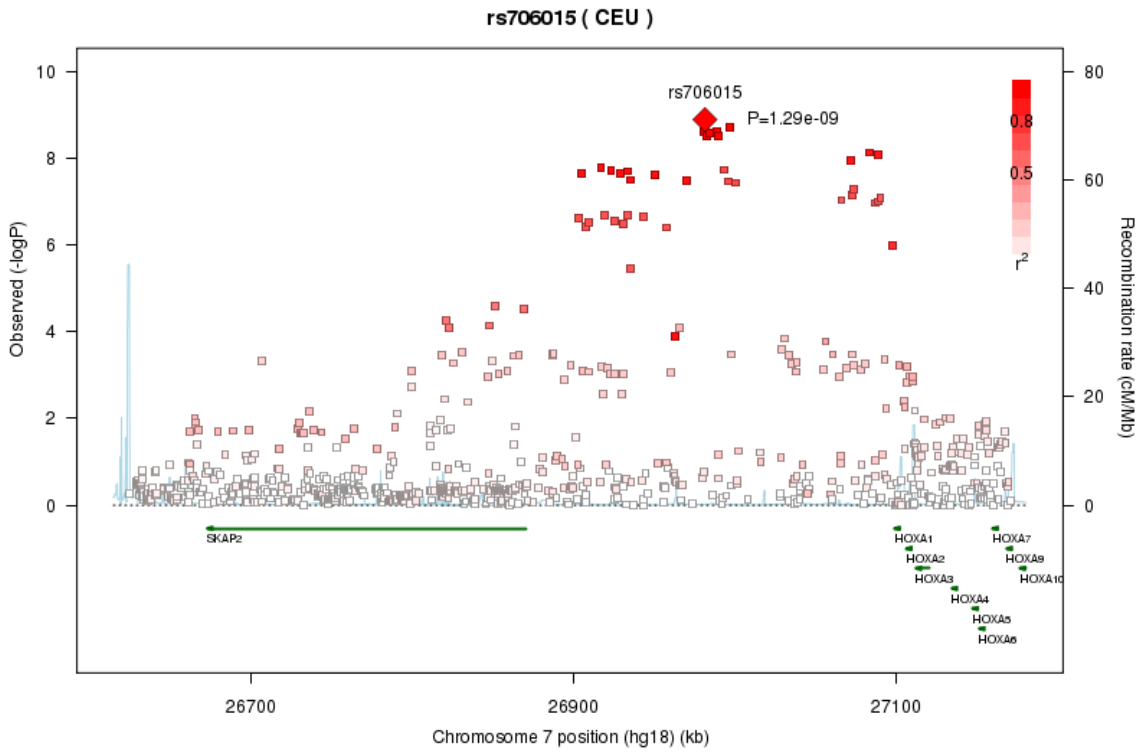
B



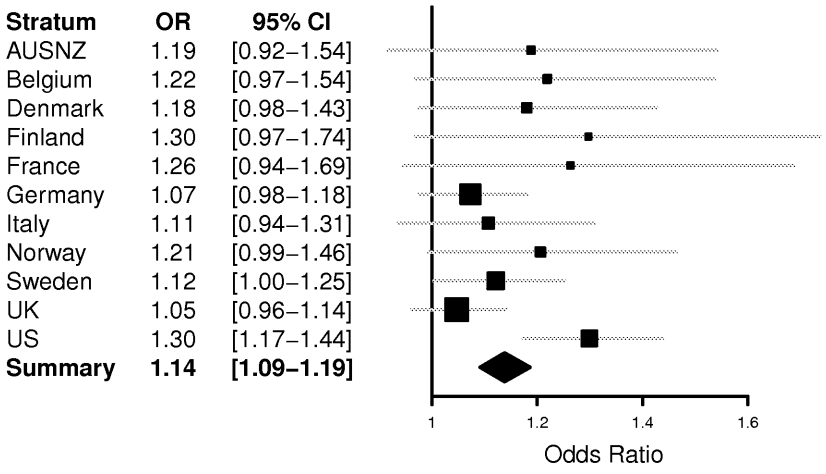
A) Regional Association and B) Forest Plot

Supplementary Figure 22. Discovery phase rs706015.

A



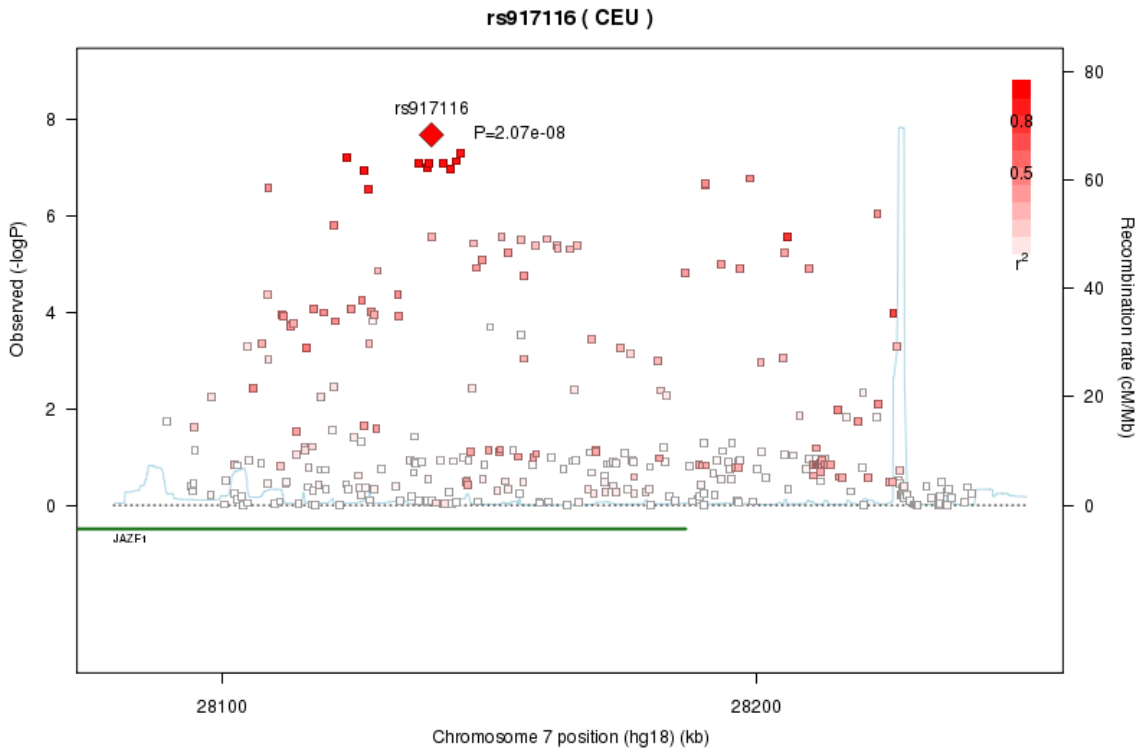
B



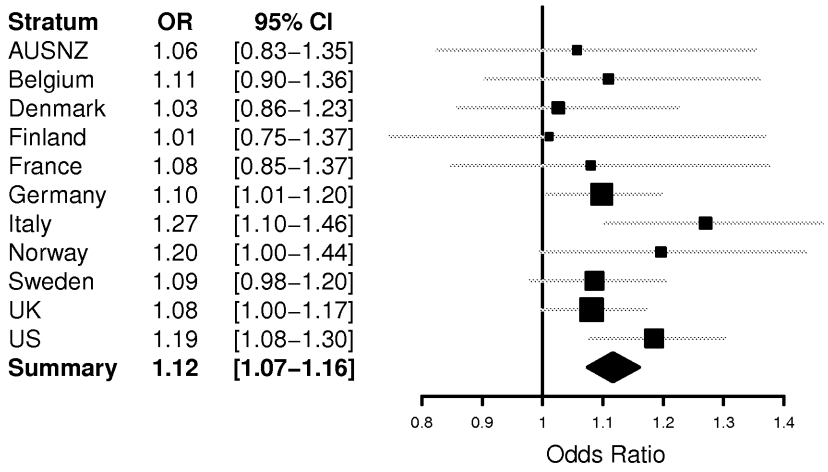
A) Regional Association and B) Forest Plot

Supplementary Figure 23. Discovery phase rs917116.

A



B

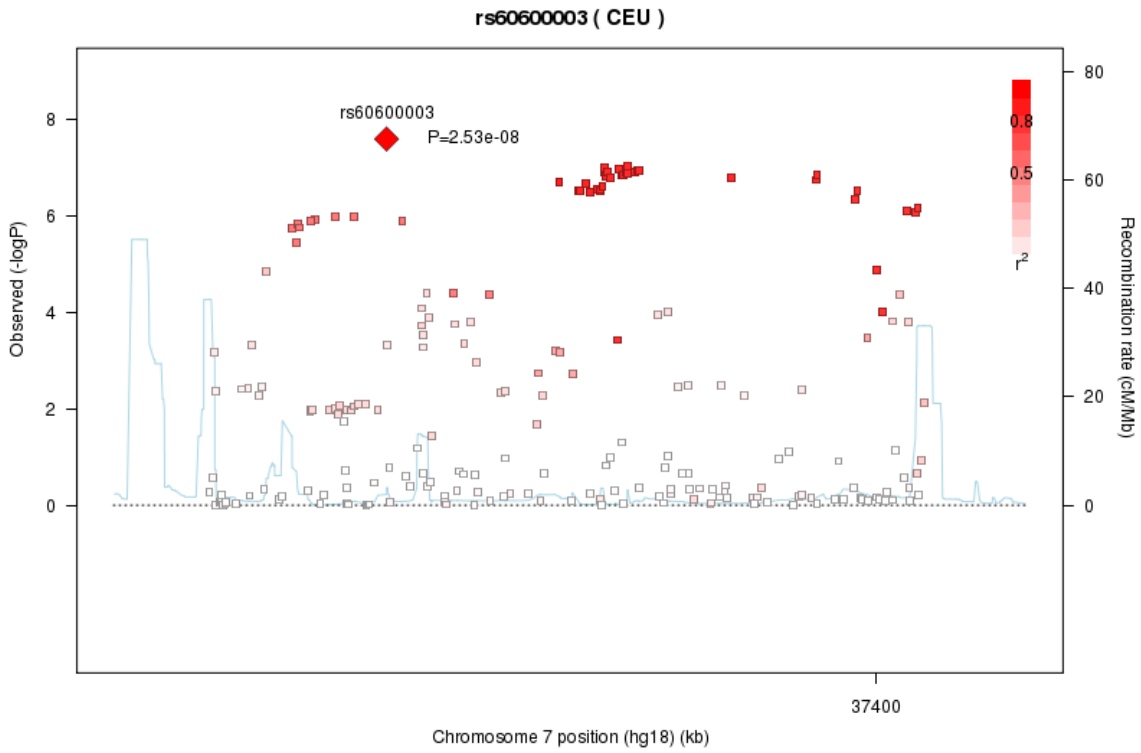


Stratum	OR	95% CI
AUSNZ	1.06	[0.83–1.35]
Belgium	1.11	[0.90–1.36]
Denmark	1.03	[0.86–1.23]
Finland	1.01	[0.75–1.37]
France	1.08	[0.85–1.37]
Germany	1.10	[1.01–1.20]
Italy	1.27	[1.10–1.46]
Norway	1.20	[1.00–1.44]
Sweden	1.09	[0.98–1.20]
UK	1.08	[1.00–1.17]
US	1.19	[1.08–1.30]
Summary	**1.12**	**[1.07–1.16]**

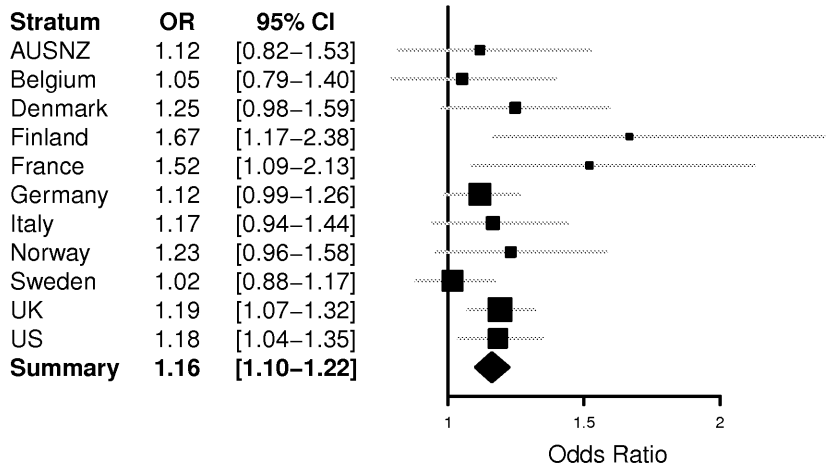
A) Regional Association and B) Forest Plot

Supplementary Figure 24. Discovery phase rs60600003.

A



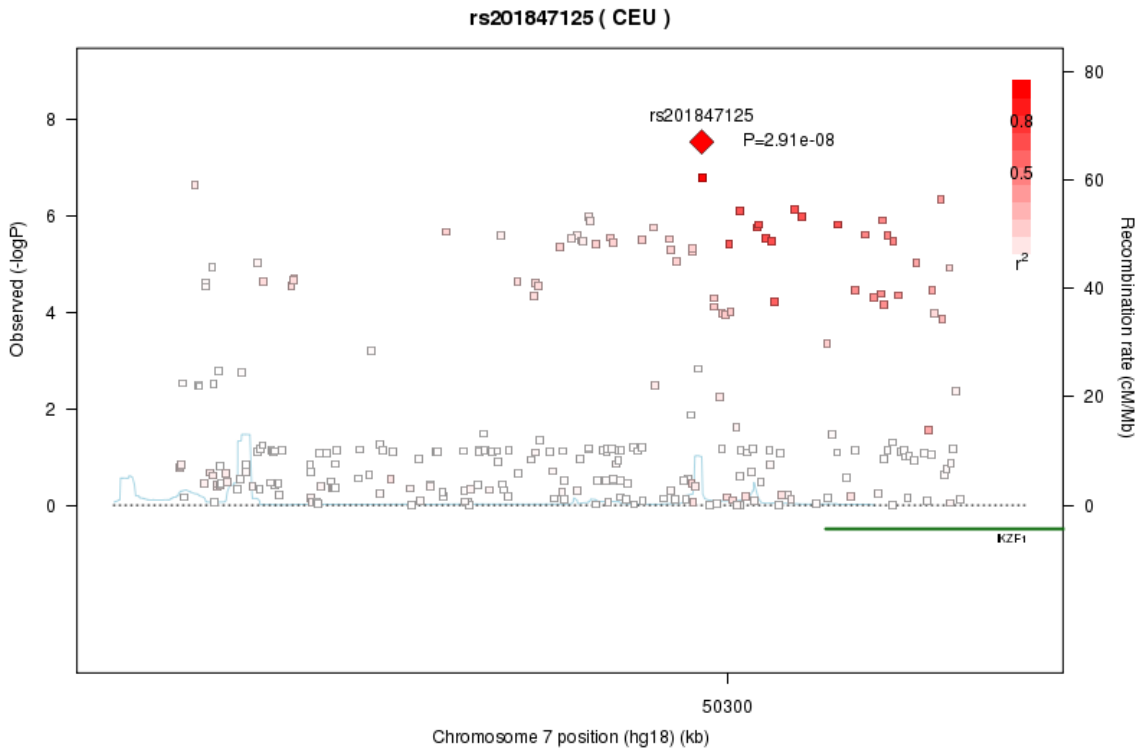
B



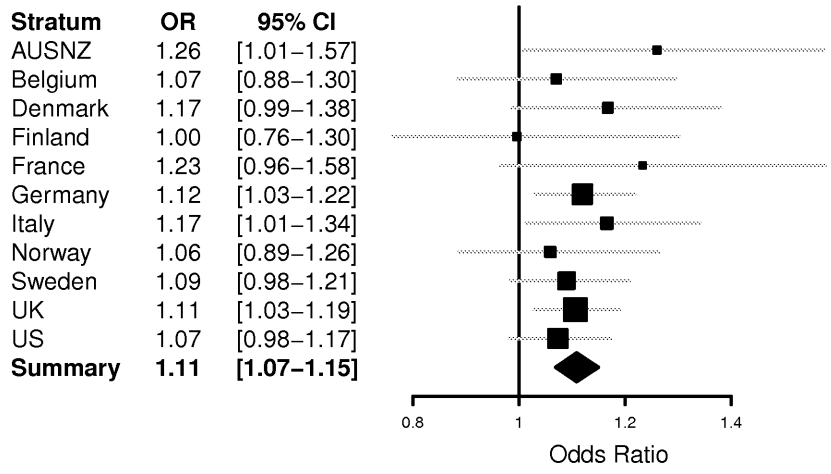
A) Regional Association and B) Forest Plot

Supplementary Figure 25. Discovery phase rs201847125.

A



B

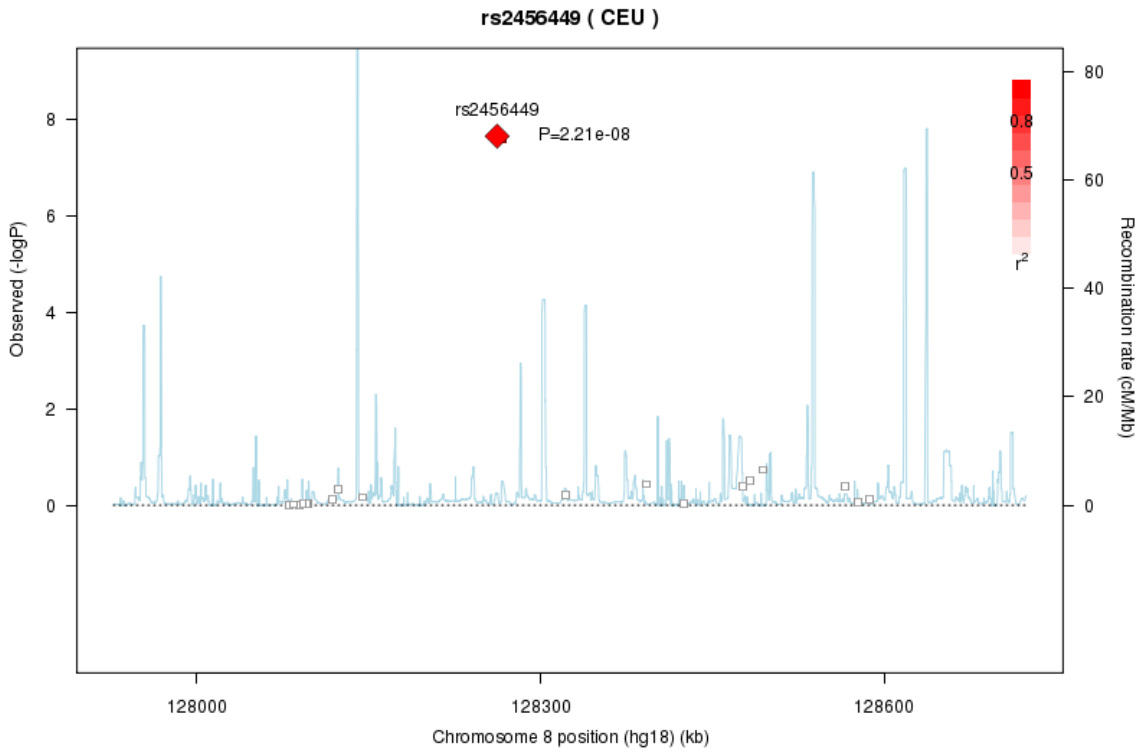


Stratum	OR	95% CI
AUSNZ	1.26	[1.01–1.57]
Belgium	1.07	[0.88–1.30]
Denmark	1.17	[0.99–1.38]
Finland	1.00	[0.76–1.30]
France	1.23	[0.96–1.58]
Germany	1.12	[1.03–1.22]
Italy	1.17	[1.01–1.34]
Norway	1.06	[0.89–1.26]
Sweden	1.09	[0.98–1.21]
UK	1.11	[1.03–1.19]
US	1.07	[0.98–1.17]
Summary	**1.11**	**[1.07–1.15]**

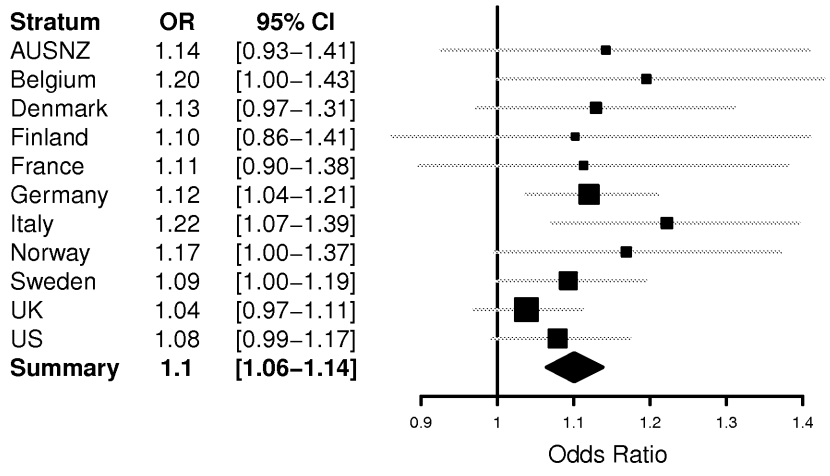
A) Regional Association and B) Forest Plot

Supplementary Figure 26. Discovery phase rs2456449.

A



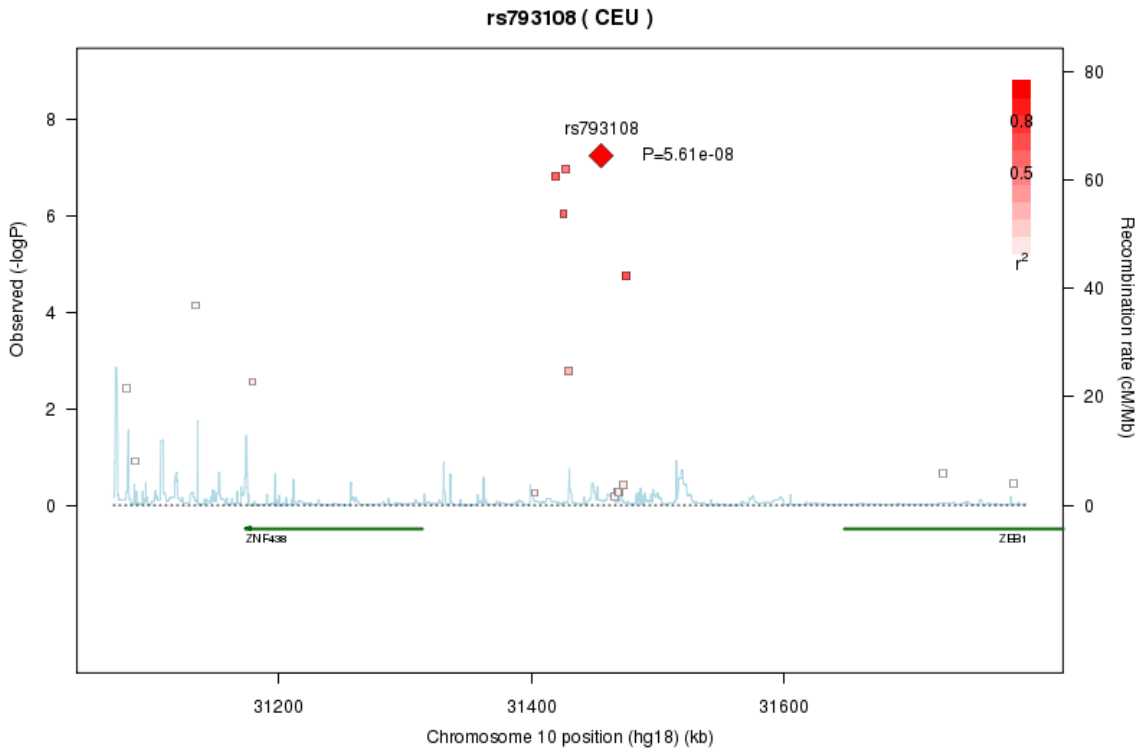
B



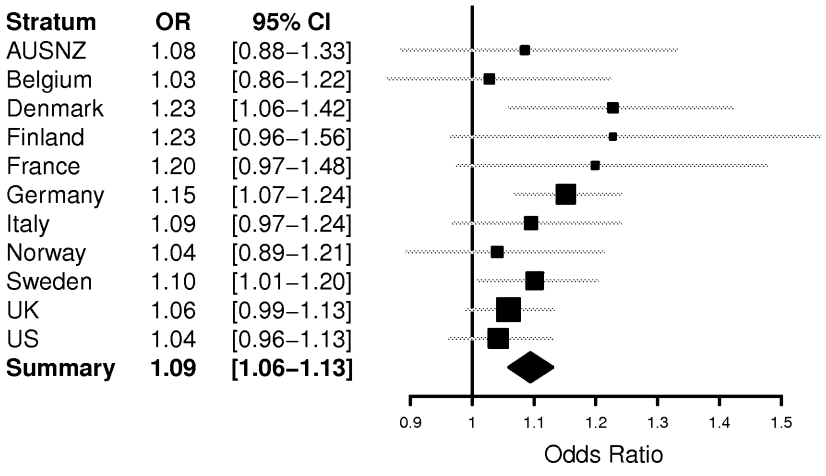
A) Regional Association and B) Forest Plot

Supplementary Figure 27. Discovery phase rs793108.

A



B

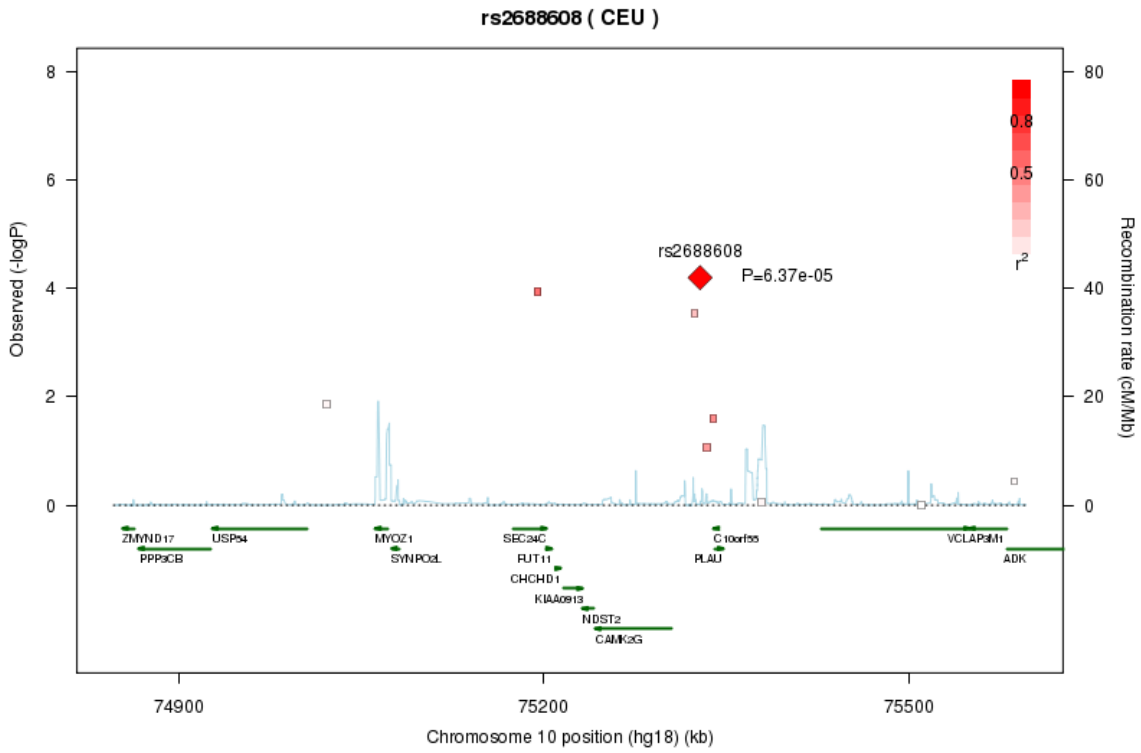


Stratum	OR	95% CI
AUSNZ	1.08	[0.88–1.33]
Belgium	1.03	[0.86–1.22]
Denmark	1.23	[1.06–1.42]
Finland	1.23	[0.96–1.56]
France	1.20	[0.97–1.48]
Germany	1.15	[1.07–1.24]
Italy	1.09	[0.97–1.24]
Norway	1.04	[0.89–1.21]
Sweden	1.10	[1.01–1.20]
UK	1.06	[0.99–1.13]
US	1.04	[0.96–1.13]
Summary	**1.09**	**[1.06–1.13]**

A) Regional Association and B) Forest Plot

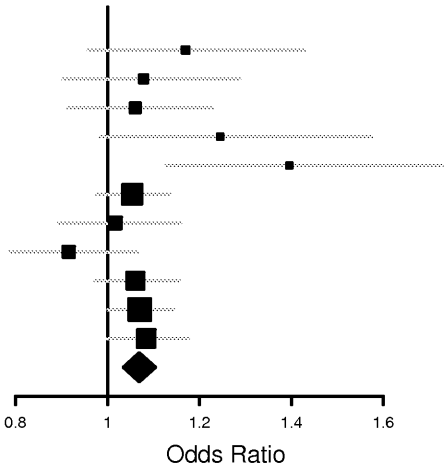
Supplementary Figure 28. Discovery phase rs2688608.

A



B

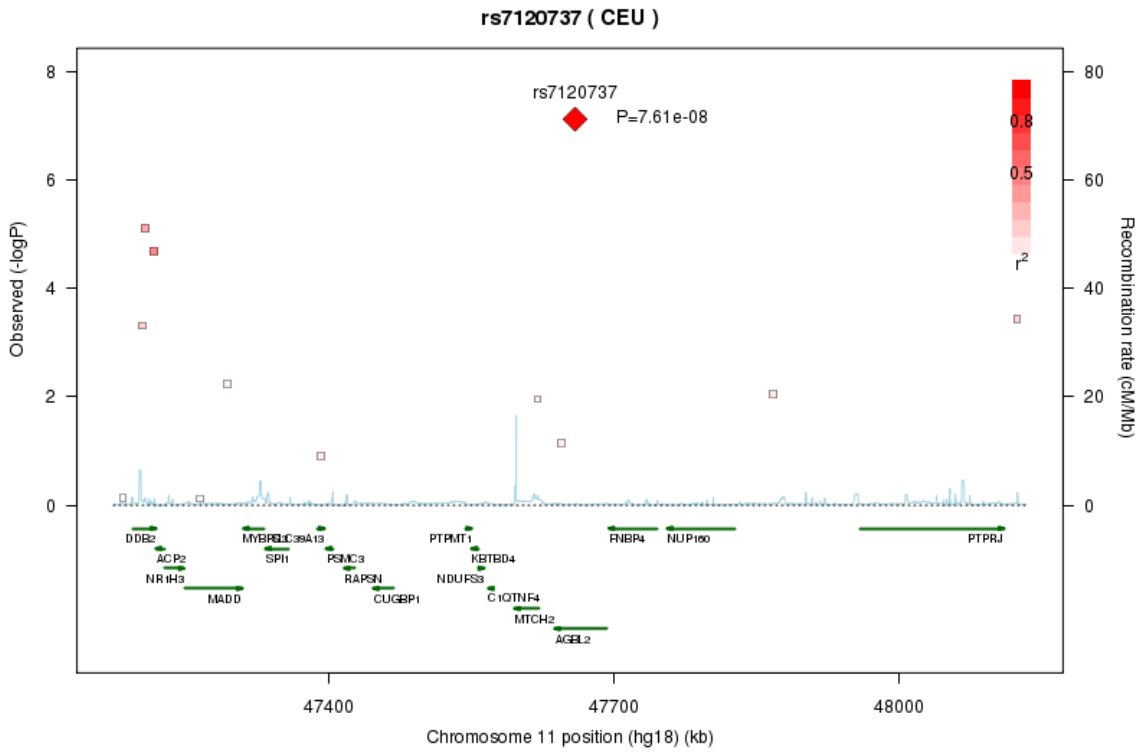
Stratum	OR	95% CI
AUSNZ	1.17	[0.96–1.43]
Belgium	1.08	[0.90–1.29]
Denmark	1.06	[0.91–1.23]
Finland	1.25	[0.98–1.58]
France	1.40	[1.13–1.73]
Germany	1.05	[0.98–1.14]
Italy	1.02	[0.89–1.16]
Norway	0.92	[0.79–1.06]
Sweden	1.06	[0.97–1.16]
UK	1.07	[1.00–1.14]
US	1.08	[1.00–1.18]
Summary	**1.07**	**[1.03–1.10]**



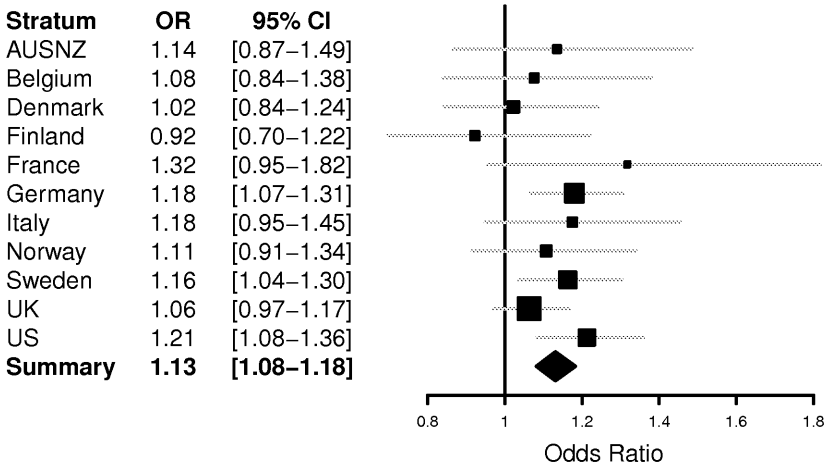
A) Regional Association and B)Forest Plot

Supplementary Figure 29. Discovery phase rs7120737.

A



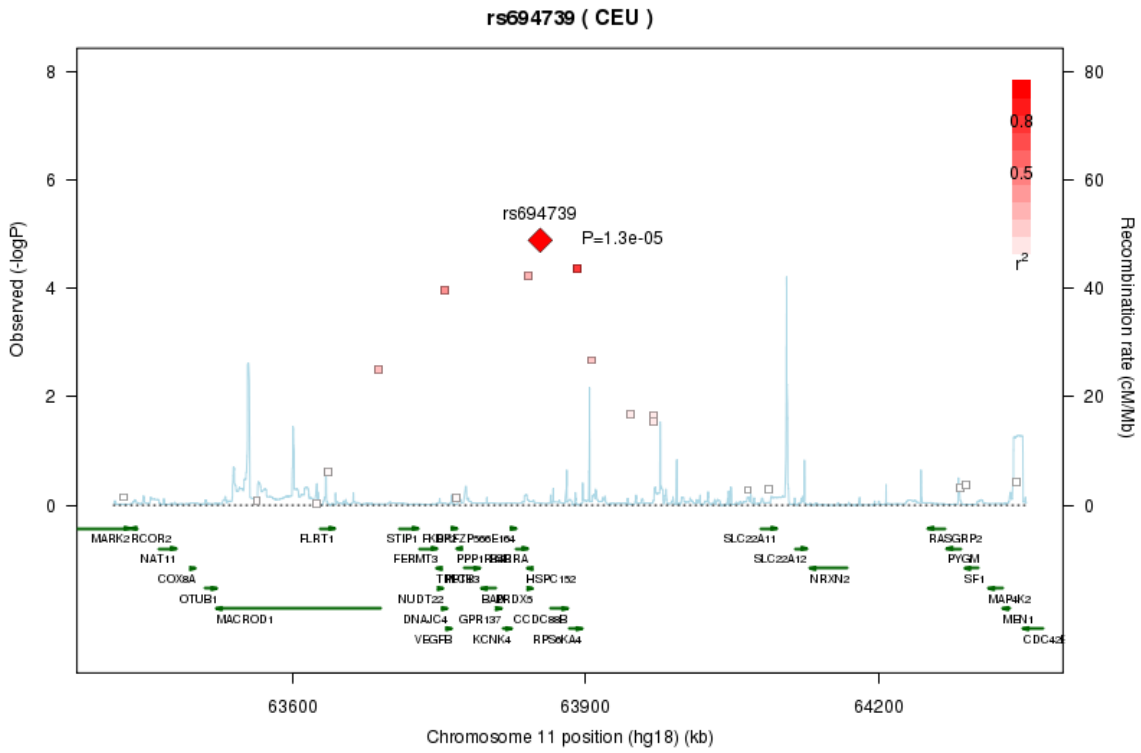
B



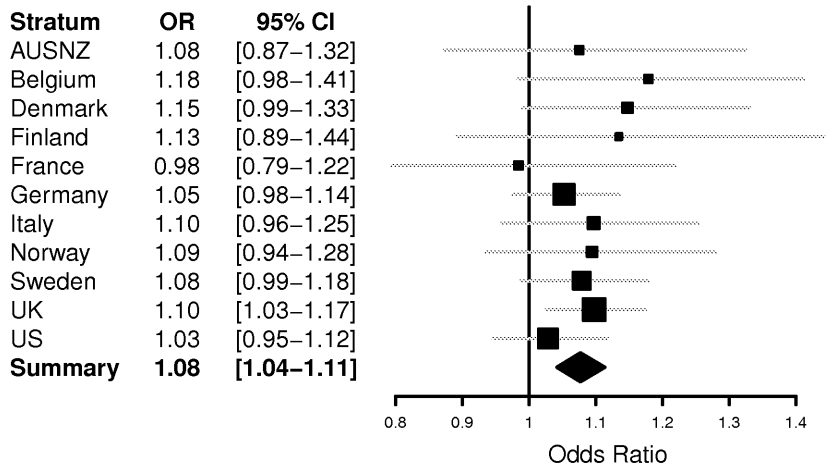
A) Regional Association and B) Forest Plot

Supplementary Figure 30. Discovery phase rs694739.

A



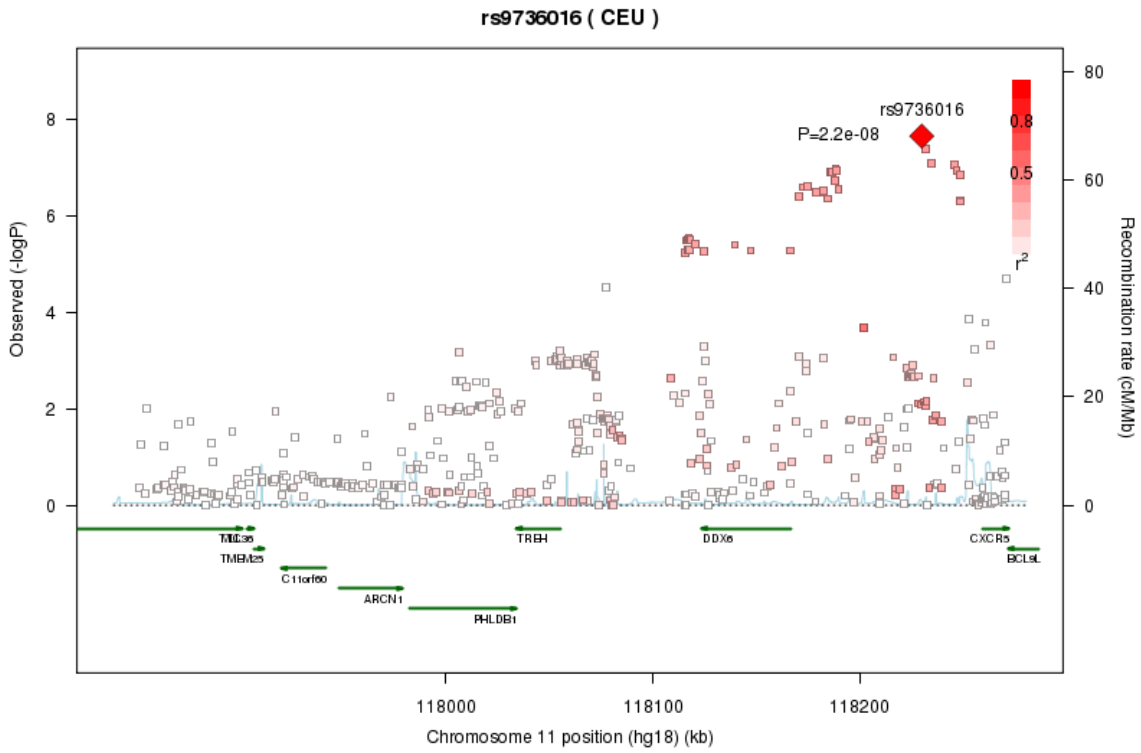
B



A) Regional Association and B) Forest Plot

Supplementary Figure 31. Discovery phase rs9736016.

A



B

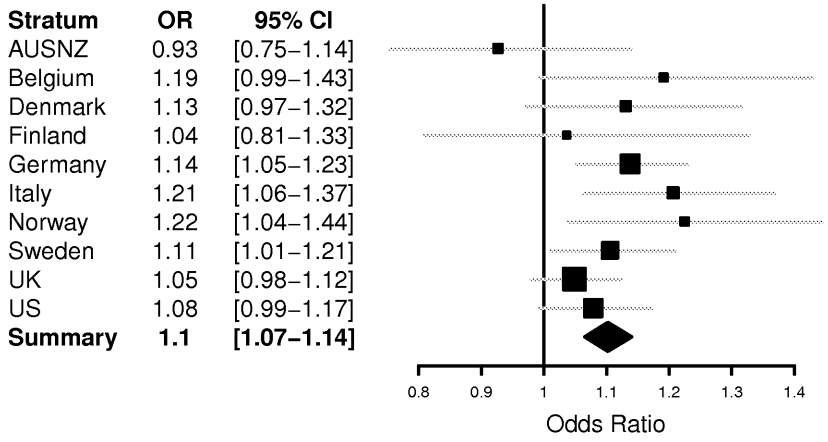
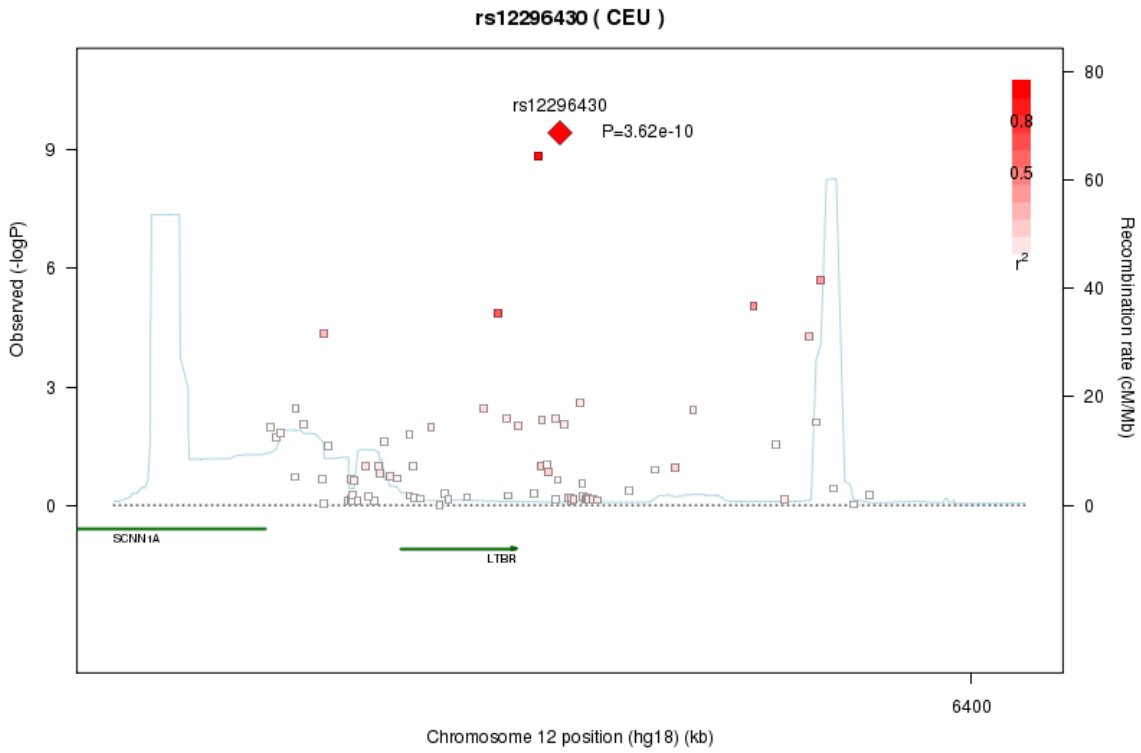


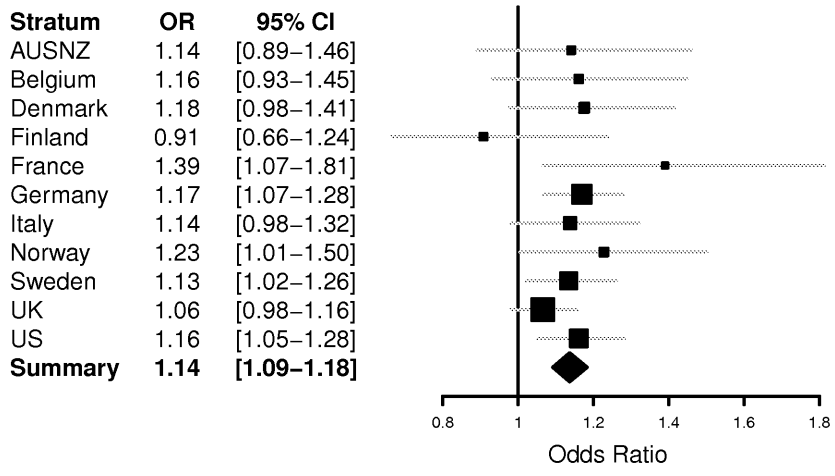
Figure S35. a) Regional Association and b) Forest Plot for rs9736016

Supplementary Figure 32. Discovery phase rs12296430.

A



B

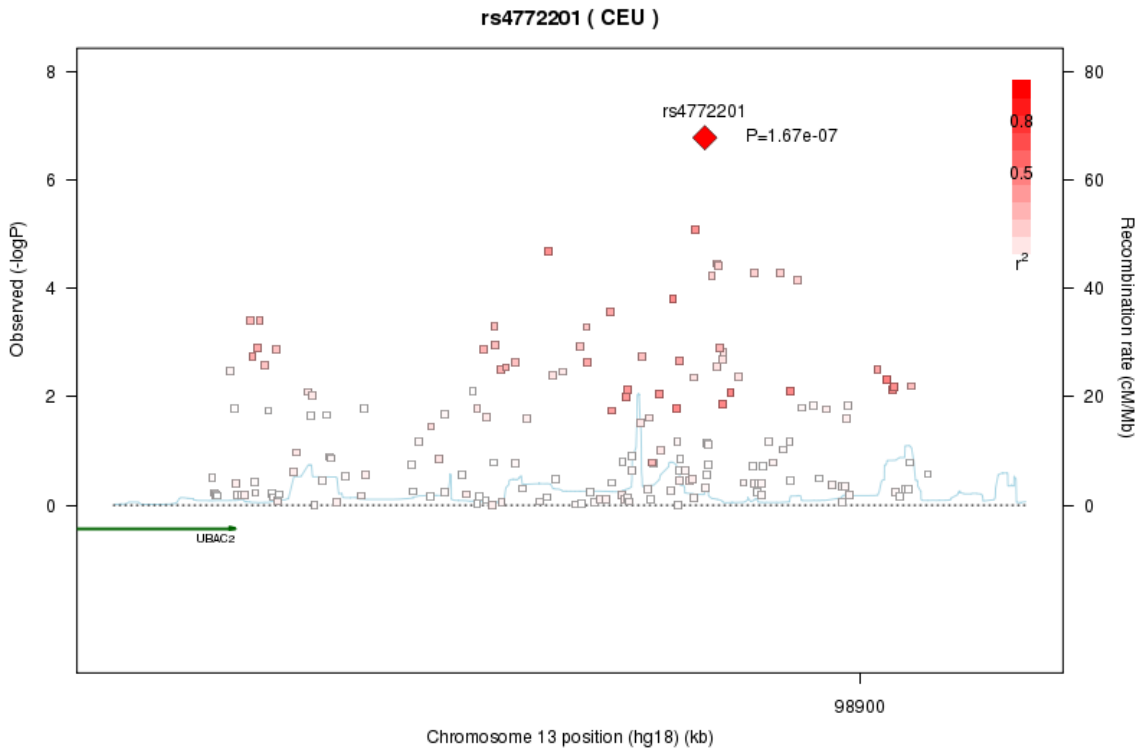


Stratum	OR	95% CI
AUSNZ	1.14	[0.89–1.46]
Belgium	1.16	[0.93–1.45]
Denmark	1.18	[0.98–1.41]
Finland	0.91	[0.66–1.24]
France	1.39	[1.07–1.81]
Germany	1.17	[1.07–1.28]
Italy	1.14	[0.98–1.32]
Norway	1.23	[1.01–1.50]
Sweden	1.13	[1.02–1.26]
UK	1.06	[0.98–1.16]
US	1.16	[1.05–1.28]
Summary	**1.14**	**[1.09–1.18]**

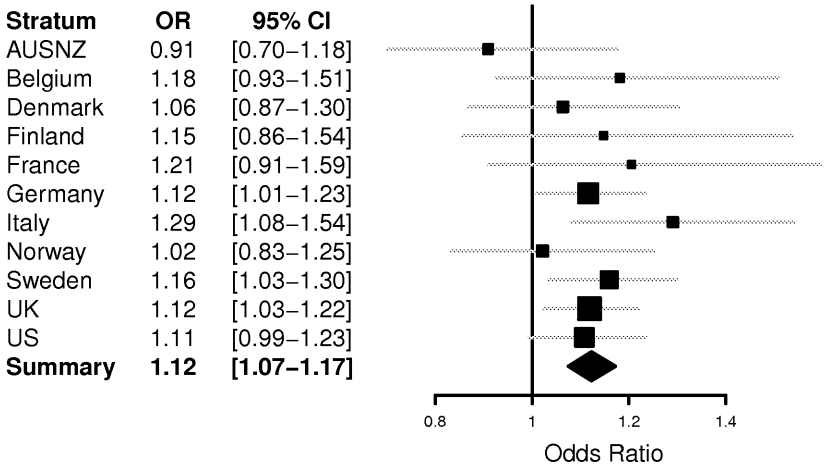
A) Regional Association and B) Forest Plot

Supplementary Figure 33. Discovery phase rs4772201.

A



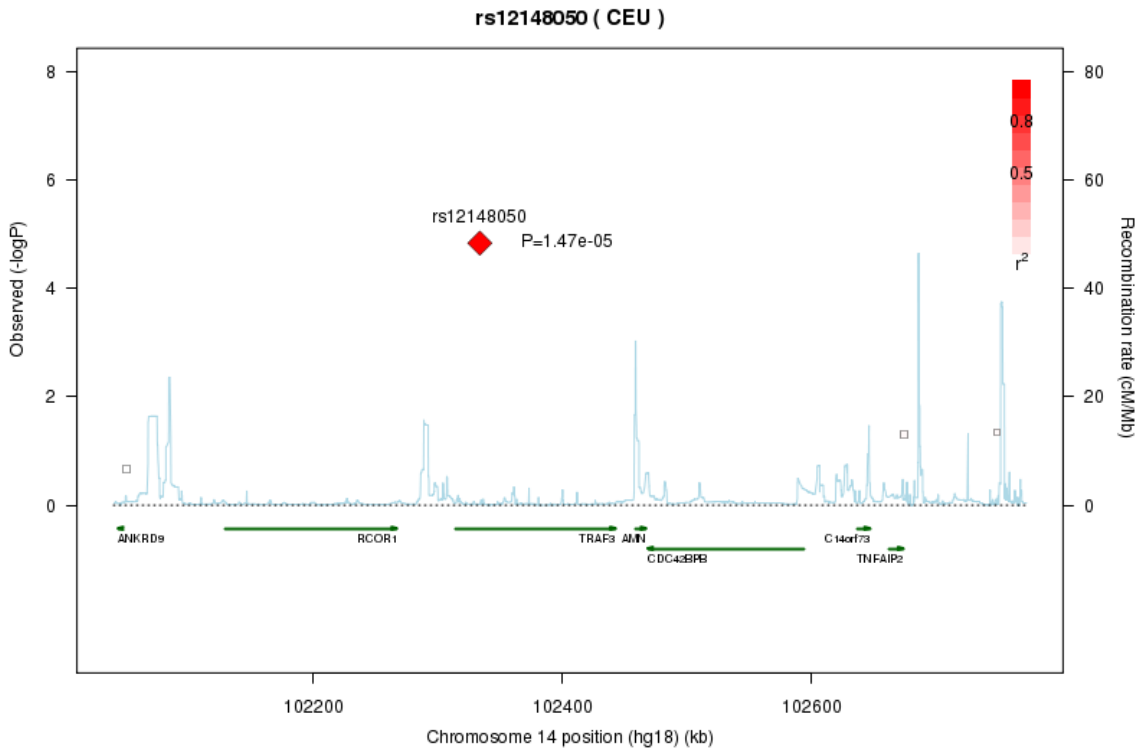
B



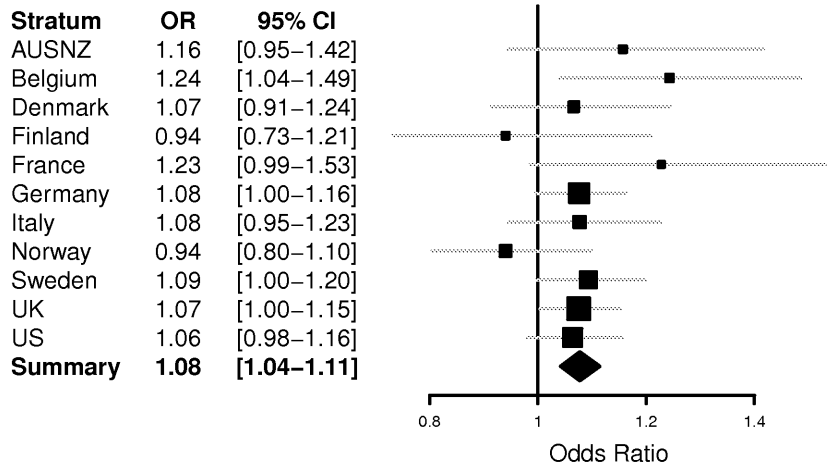
A) Regional Association and B) Forest Plot for rs4772201

Supplementary Figure 34. Discovery phase rs12148050.

A



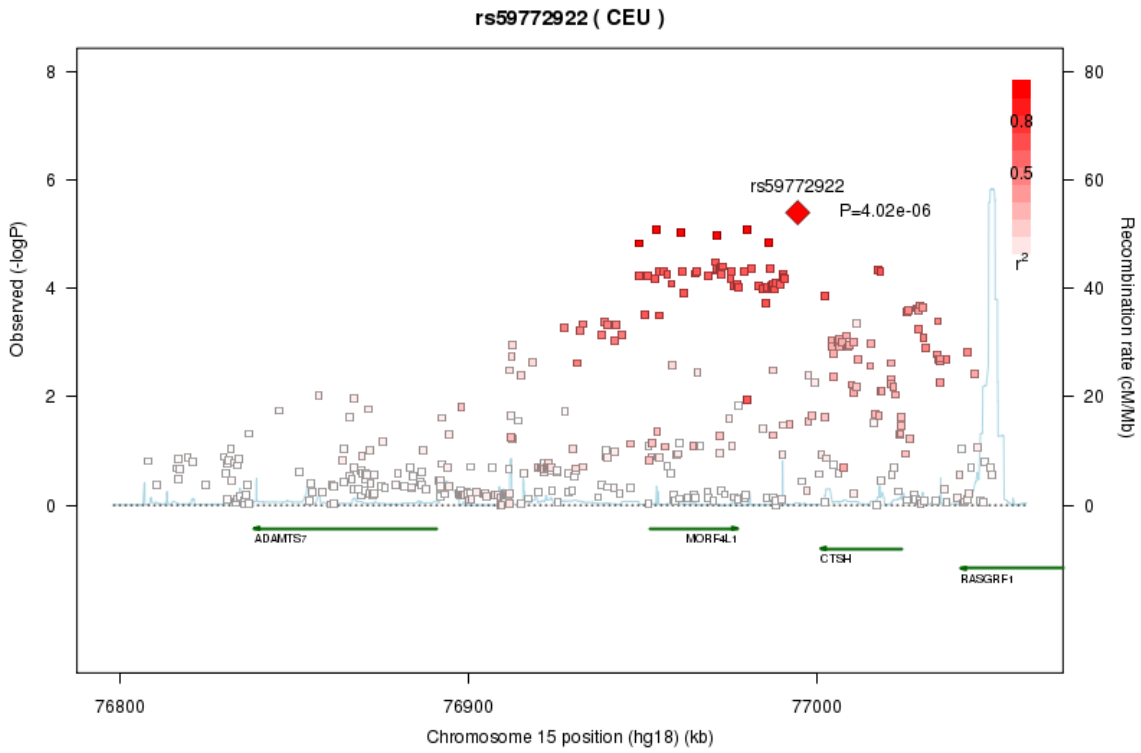
B



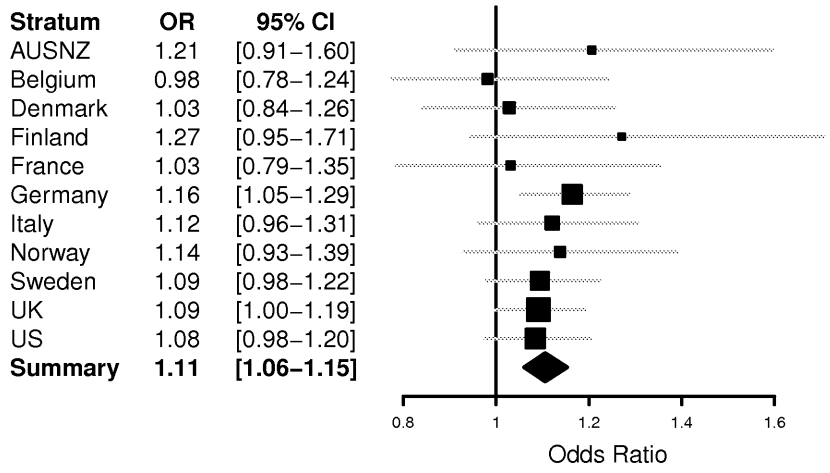
A) Regional Association and B) Forest Plot

Supplementary Figure 35. Discovery phase rs59772922.

A



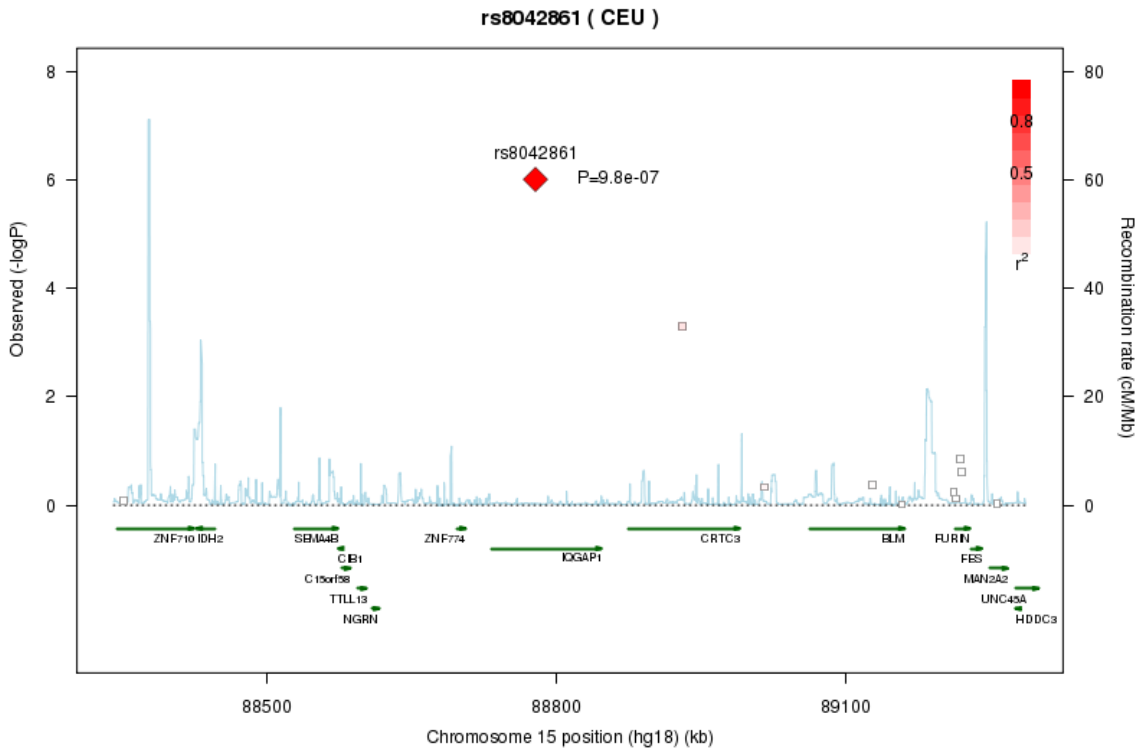
B



A) Regional Association and B) Forest Plot

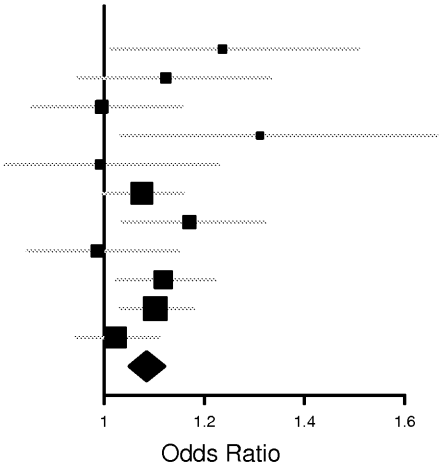
Supplementary Figure 36. Discovery phase rs8042861.

A



B

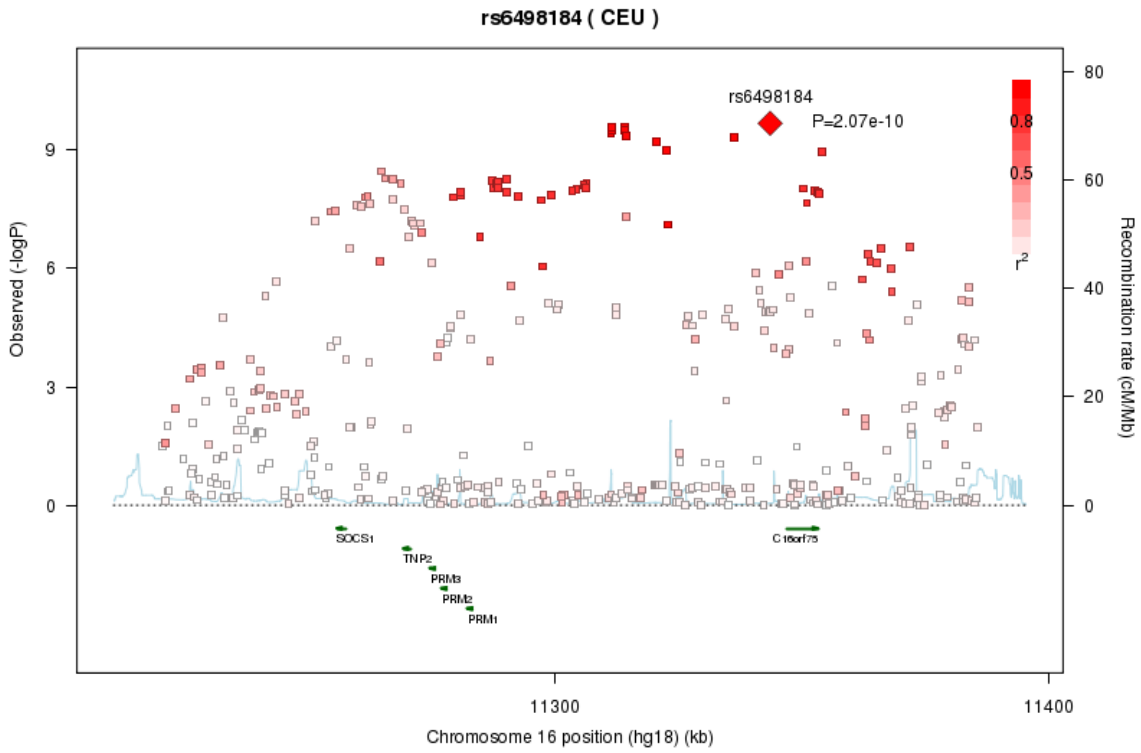
Stratum	OR	95% CI
AUSNZ	1.24	[1.01–1.51]
Belgium	1.12	[0.95–1.33]
Denmark	0.99	[0.86–1.16]
Finland	1.31	[1.03–1.66]
France	0.99	[0.80–1.23]
Germany	1.07	[1.00–1.16]
Italy	1.17	[1.04–1.32]
Norway	0.99	[0.85–1.15]
Sweden	1.12	[1.02–1.22]
UK	1.10	[1.03–1.18]
US	1.02	[0.94–1.11]
Summary	**1.08**	**[1.05–1.12]**



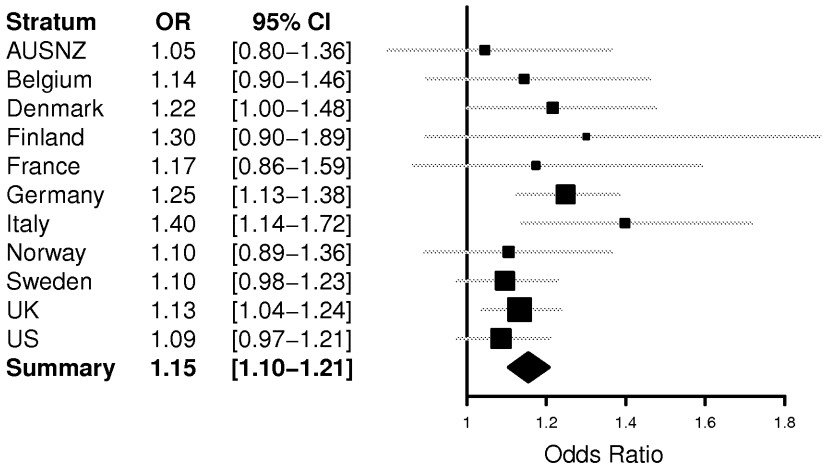
A) Regional Association and B) Forest Plot

Supplementary Figure 37. Discovery phase rs6498184.

A



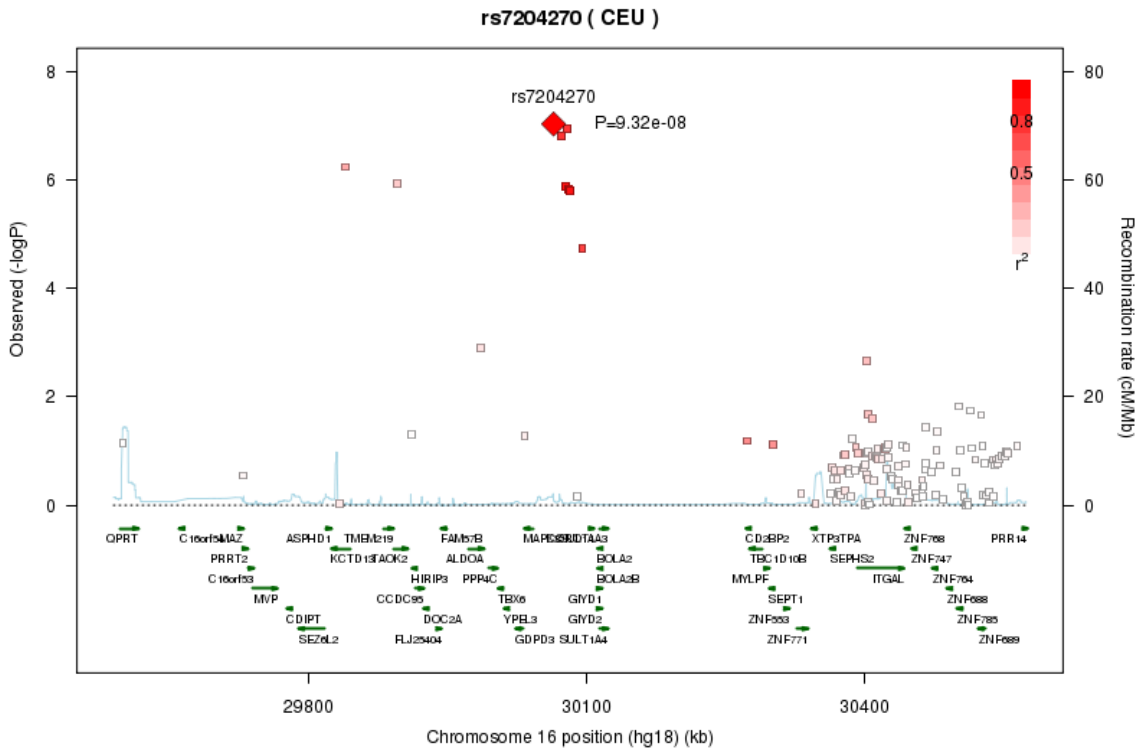
B



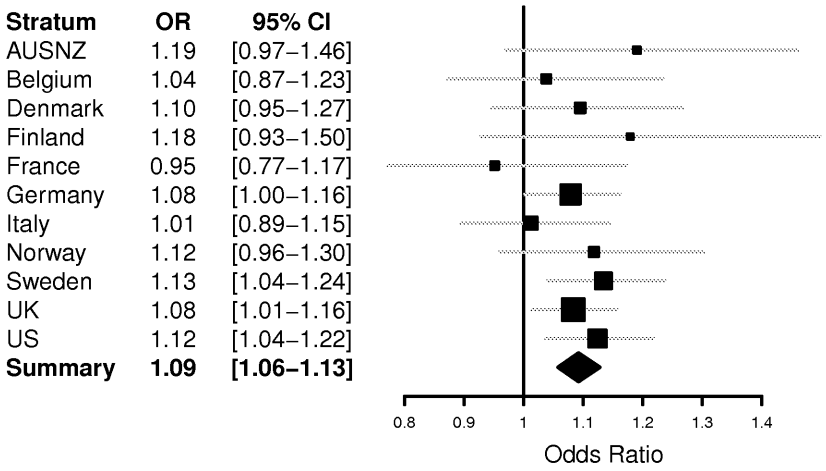
A) Regional Association and B) Forest Plot

Supplementary Figure 38. Discovery phase rs7204270.

A



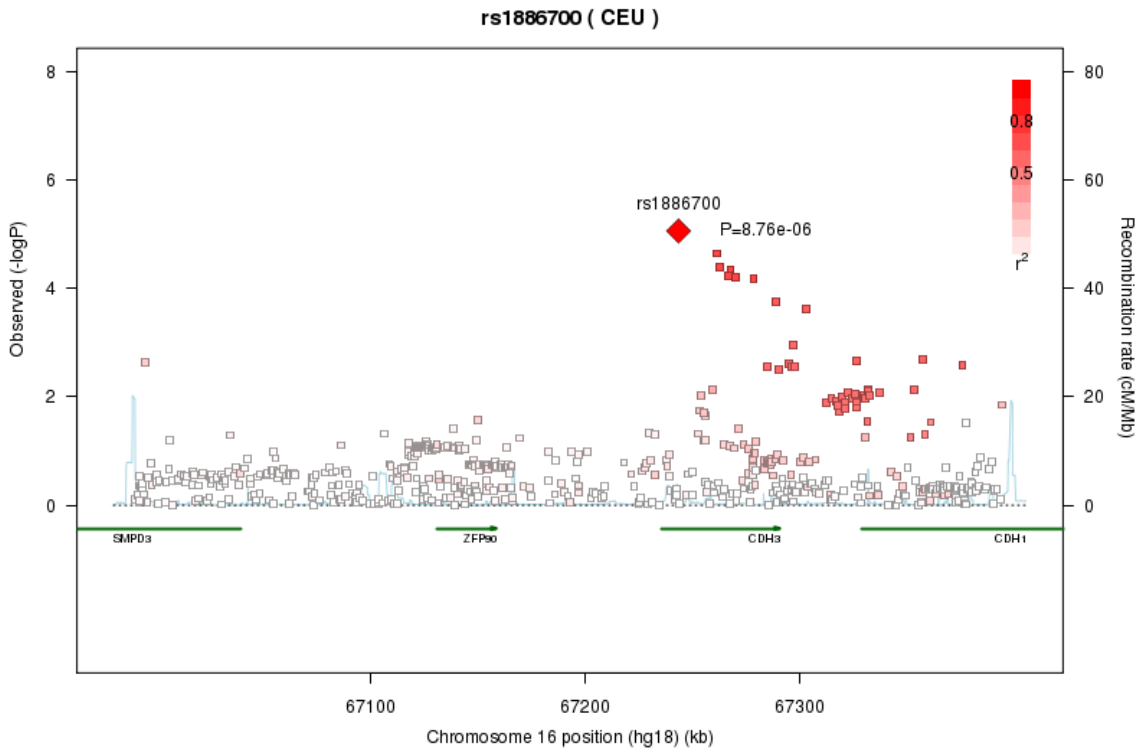
B



A) Regional Association and B) Forest Plot

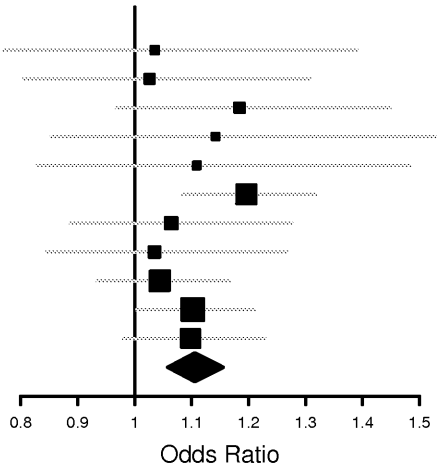
Supplementary Figure 39. Discovery phase rs1886700.

A



B

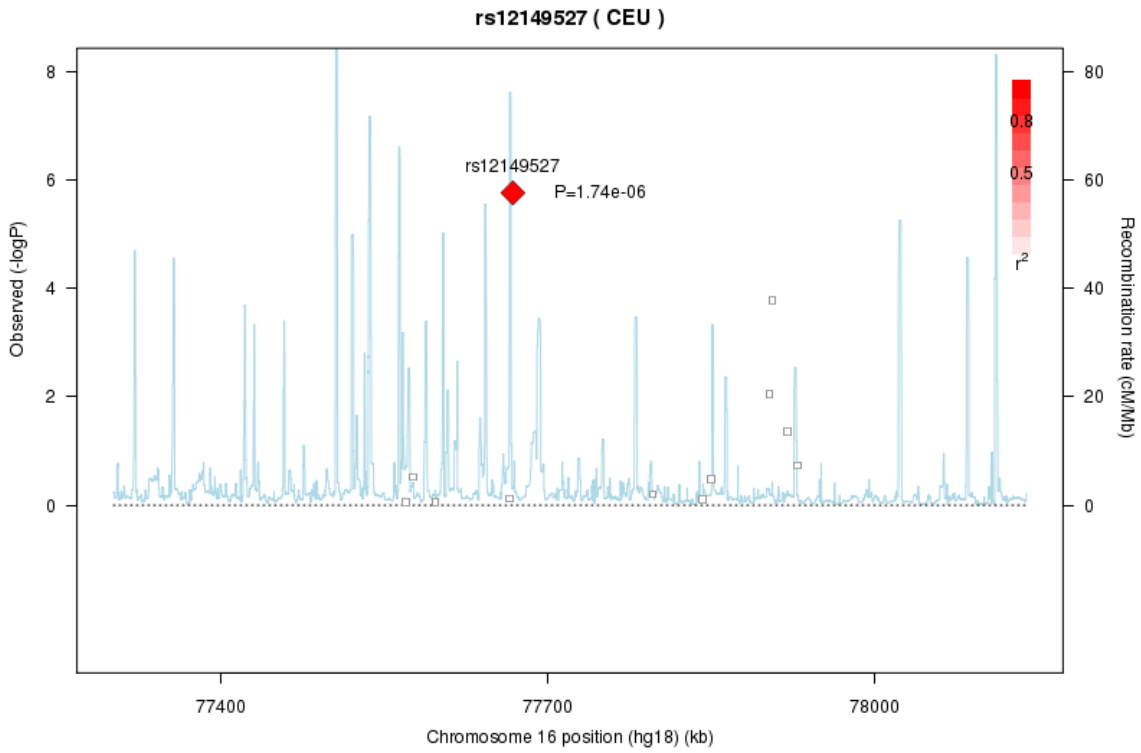
Stratum	OR	95% CI
AUSNZ	1.03	[0.77–1.39]
Belgium	1.03	[0.80–1.31]
Denmark	1.18	[0.97–1.45]
Finland	1.14	[0.85–1.53]
France	1.11	[0.83–1.48]
Germany	1.20	[1.08–1.32]
Italy	1.06	[0.89–1.28]
Norway	1.03	[0.85–1.27]
Sweden	1.04	[0.93–1.17]
UK	1.10	[1.00–1.21]
US	1.10	[0.98–1.23]
Summary	**1.11**	**[1.06–1.16]**



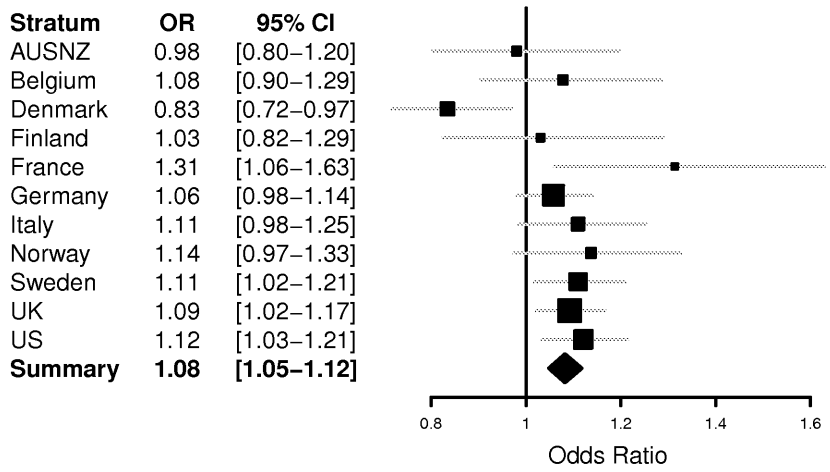
A) Regional Association and B) Forest Plot

Supplementary Figure 40. Discovery phase rs12149527.

A



B

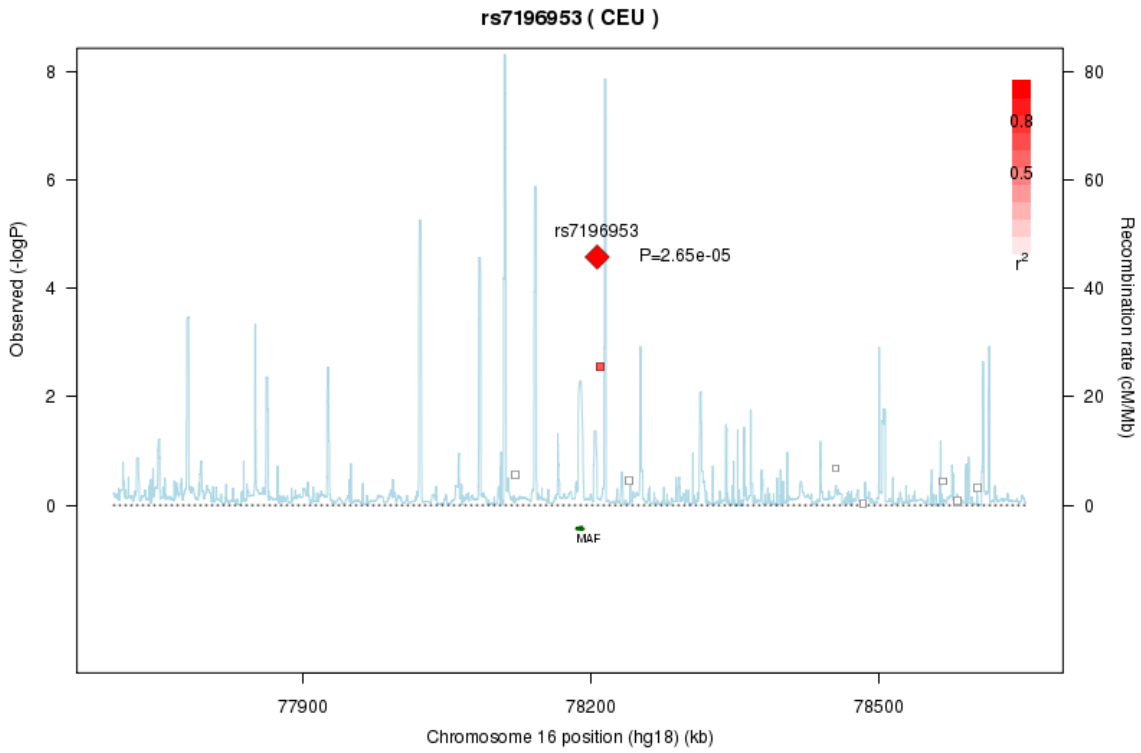


Stratum	OR	95% CI
AUSNZ	0.98	[0.80–1.20]
Belgium	1.08	[0.90–1.29]
Denmark	0.83	[0.72–0.97]
Finland	1.03	[0.82–1.29]
France	1.31	[1.06–1.63]
Germany	1.06	[0.98–1.14]
Italy	1.11	[0.98–1.25]
Norway	1.14	[0.97–1.33]
Sweden	1.11	[1.02–1.21]
UK	1.09	[1.02–1.17]
US	1.12	[1.03–1.21]
Summary	**1.08**	**[1.05–1.12]**

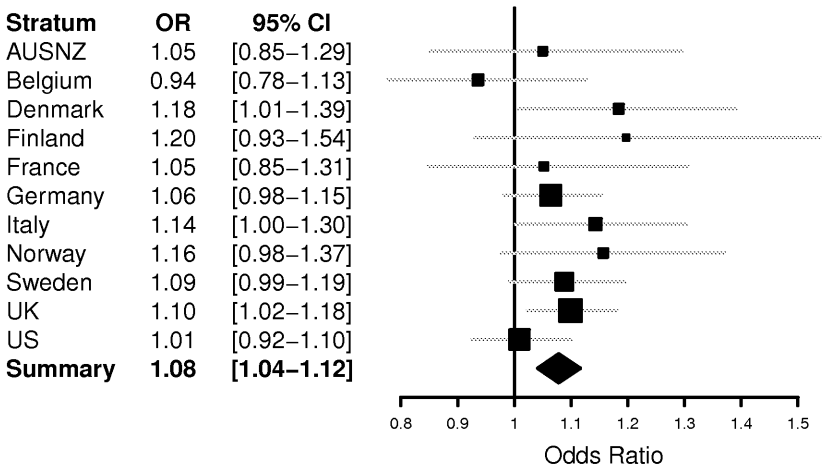
A) Regional Association and B) Forest Plot

Supplementary Figure 41. Discovery phase rs7196953.

A



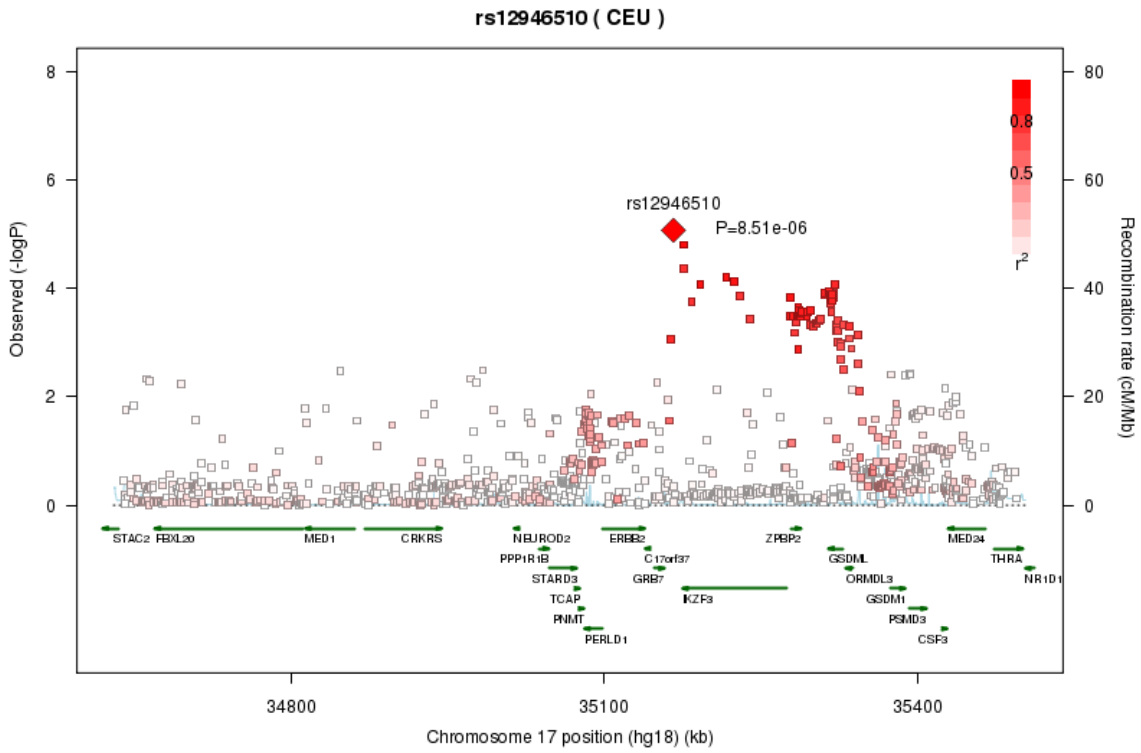
B



A) Regional Association and B) Forest Plot

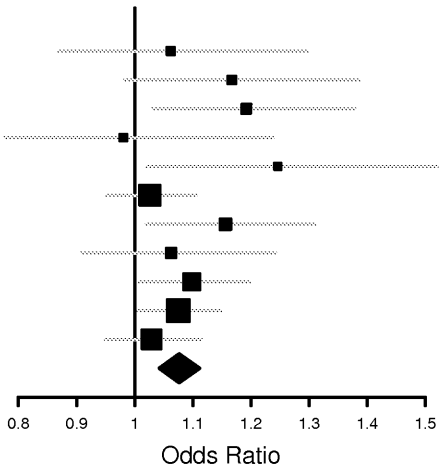
Supplementary Figure 42. Discovery phase rs12946510.

A



B

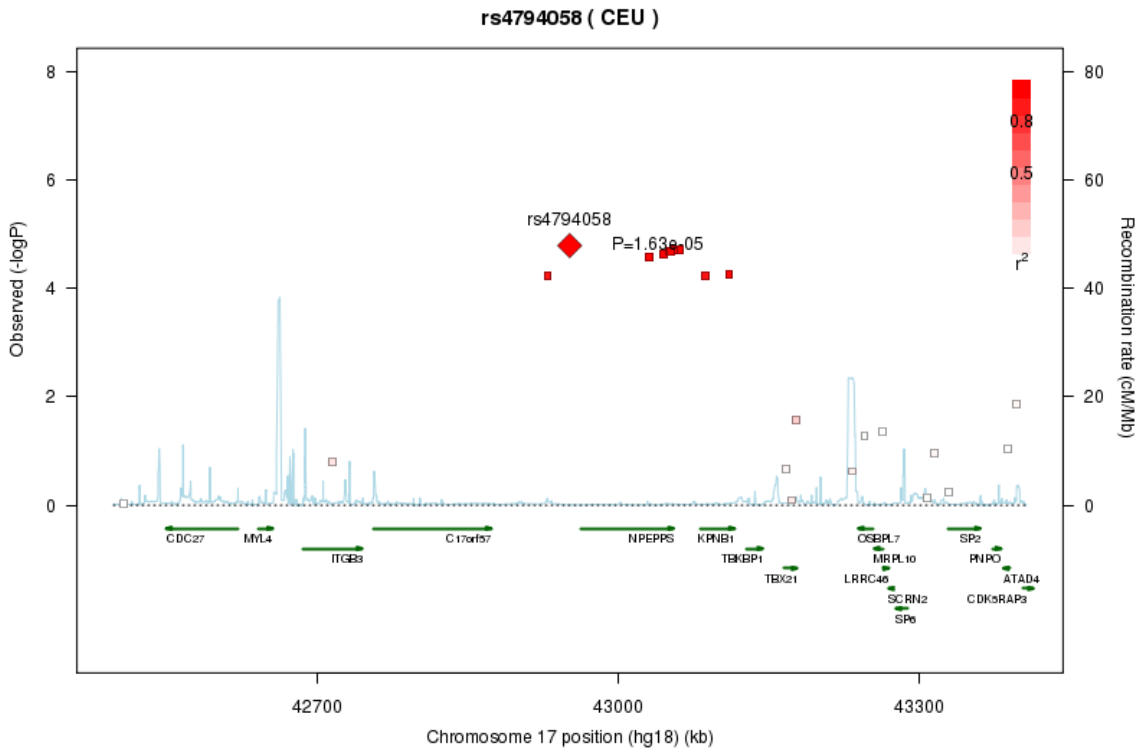
Stratum	OR	95% CI
AUSNZ	1.06	[0.87–1.30]
Belgium	1.17	[0.98–1.39]
Denmark	1.19	[1.03–1.38]
Finland	0.98	[0.78–1.24]
France	1.25	[1.02–1.52]
Germany	1.03	[0.95–1.11]
Italy	1.16	[1.02–1.31]
Norway	1.06	[0.91–1.24]
Sweden	1.10	[1.01–1.20]
UK	1.07	[1.01–1.15]
US	1.03	[0.95–1.11]
Summary	**1.08**	**[1.04–1.11]**



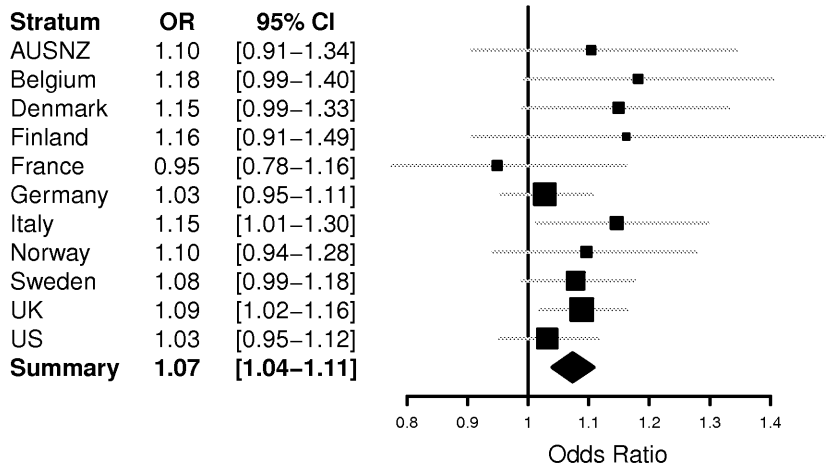
A) Regional Association and B) Forest Plot

Supplementary Figure 43. Discovery phase rs4794058.

A



B

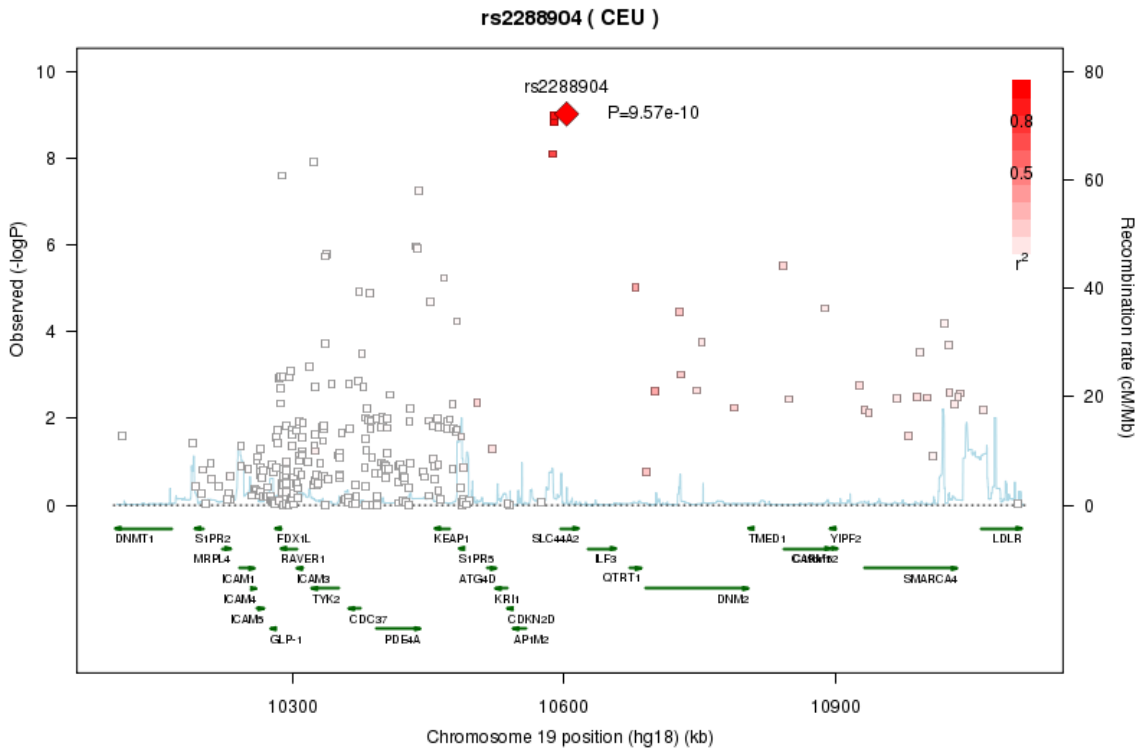


Stratum	OR	95% CI
AUSNZ	1.10	[0.91–1.34]
Belgium	1.18	[0.99–1.40]
Denmark	1.15	[0.99–1.33]
Finland	1.16	[0.91–1.49]
France	0.95	[0.78–1.16]
Germany	1.03	[0.95–1.11]
Italy	1.15	[1.01–1.30]
Norway	1.10	[0.94–1.28]
Sweden	1.08	[0.99–1.18]
UK	1.09	[1.02–1.16]
US	1.03	[0.95–1.12]
Summary	**1.07**	**[1.04–1.11]**

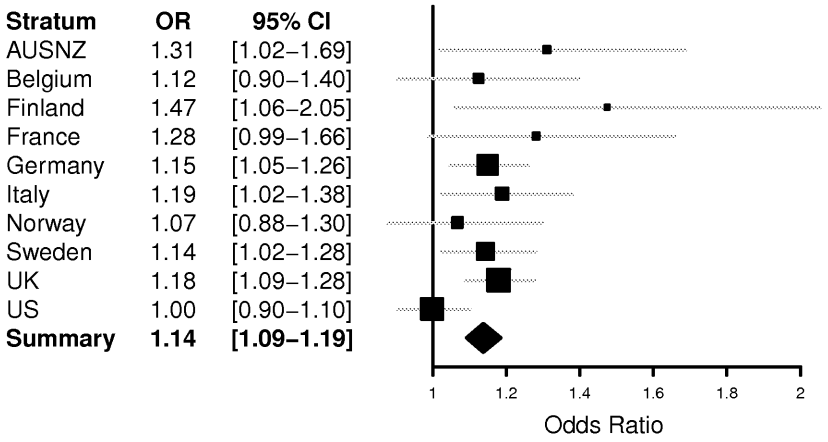
A) Regional Association and B) Forest Plot

Supplementary Figure 44. Discovery phase rs2288904.

A



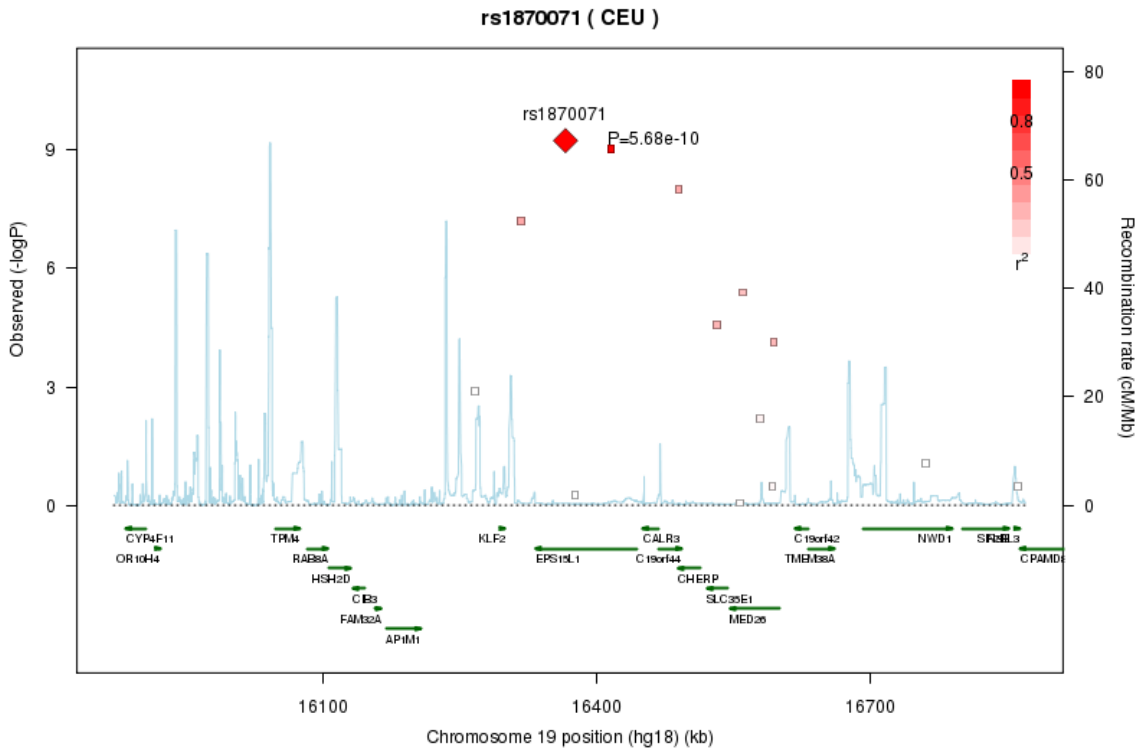
B



A) Regional Association and B) Forest Plot

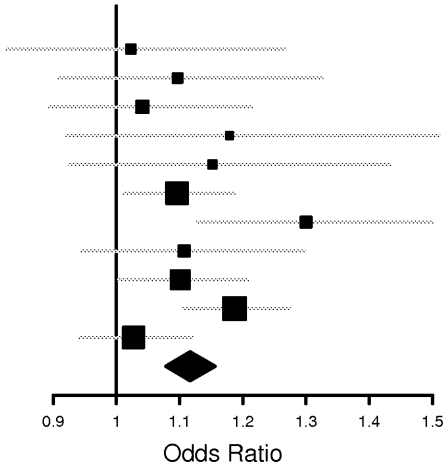
Supplementary Figure 45. Discovery phase rs1870071.

A



B

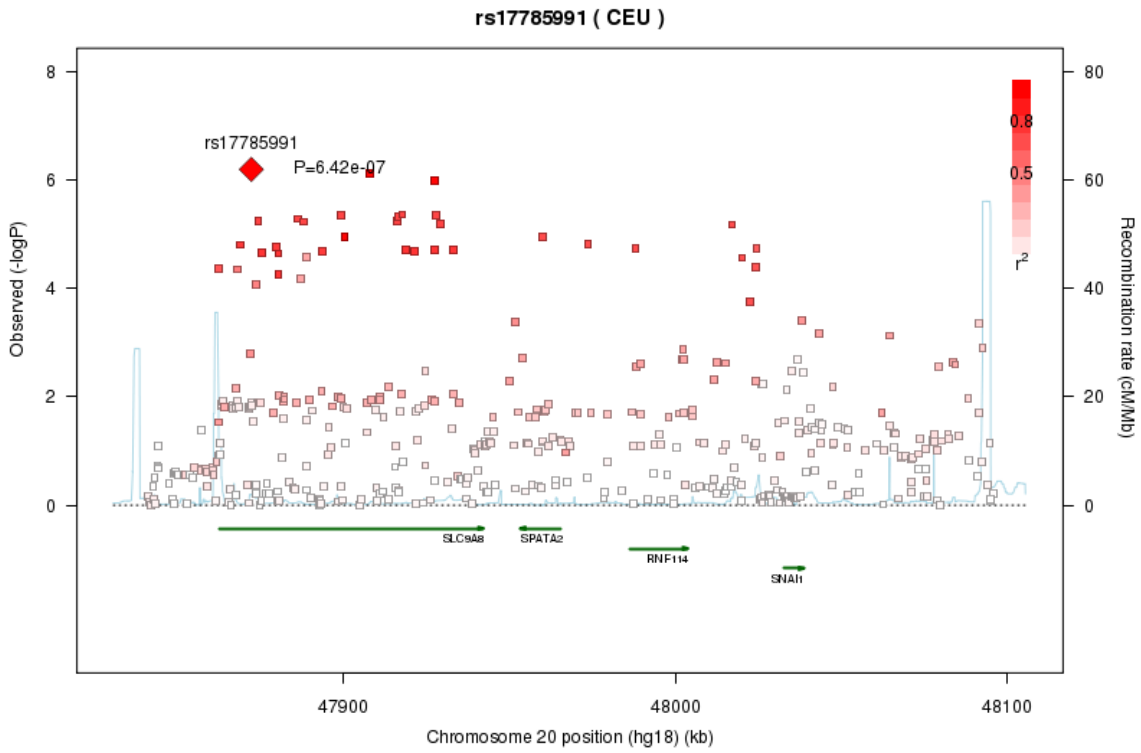
Stratum	OR	95% CI
AUSNZ	1.02	[0.83–1.27]
Belgium	1.10	[0.91–1.33]
Denmark	1.04	[0.89–1.21]
Finland	1.18	[0.92–1.51]
France	1.15	[0.93–1.43]
Germany	1.10	[1.01–1.19]
Italy	1.30	[1.13–1.50]
Norway	1.11	[0.94–1.30]
Sweden	1.10	[1.00–1.21]
UK	1.19	[1.11–1.27]
US	1.03	[0.94–1.12]
Summary	**1.12**	**[1.08–1.16]**



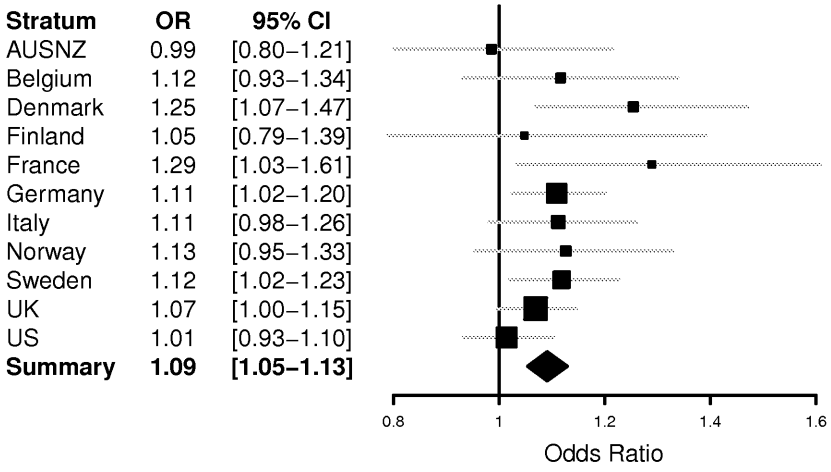
A) Regional Association and B) Forest Plot

Supplementary Figure 46. Discovery phase rs17785991.

A



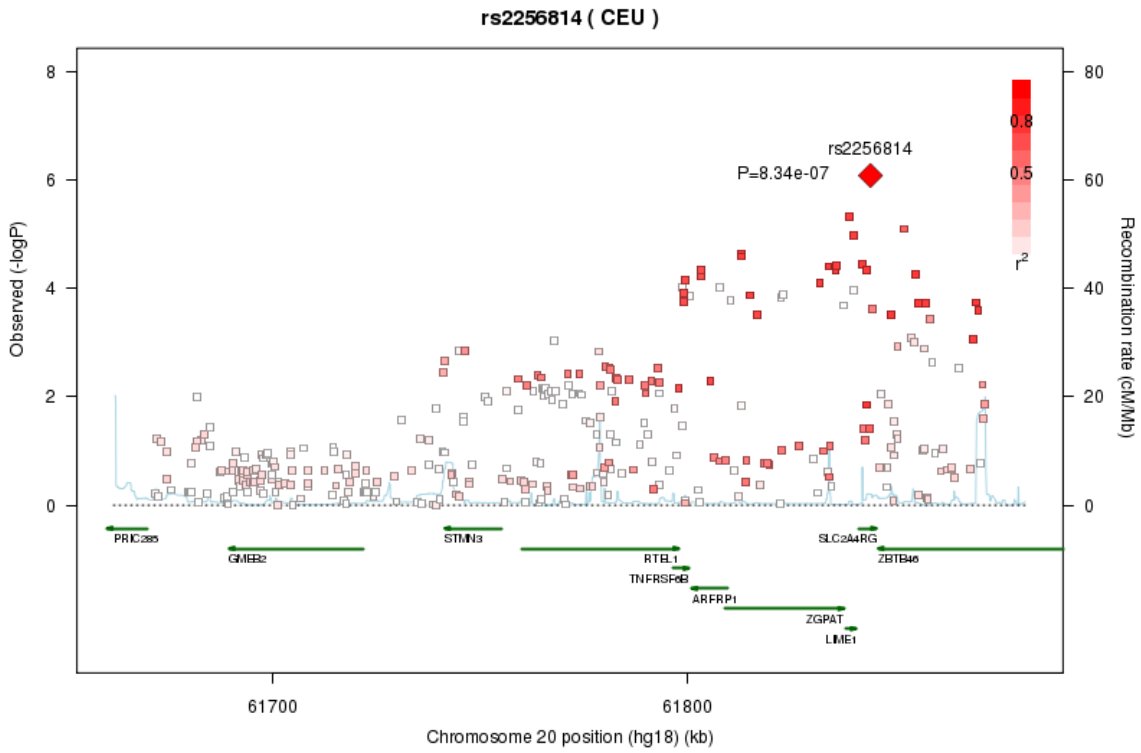
B



A) Regional Association and B) Forest Plot

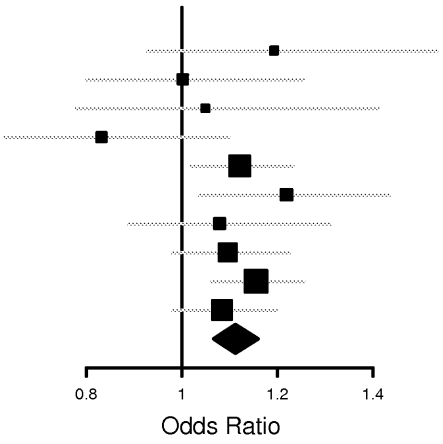
Supplementary Figure 47. Discovery phase rs2256814.

A



B

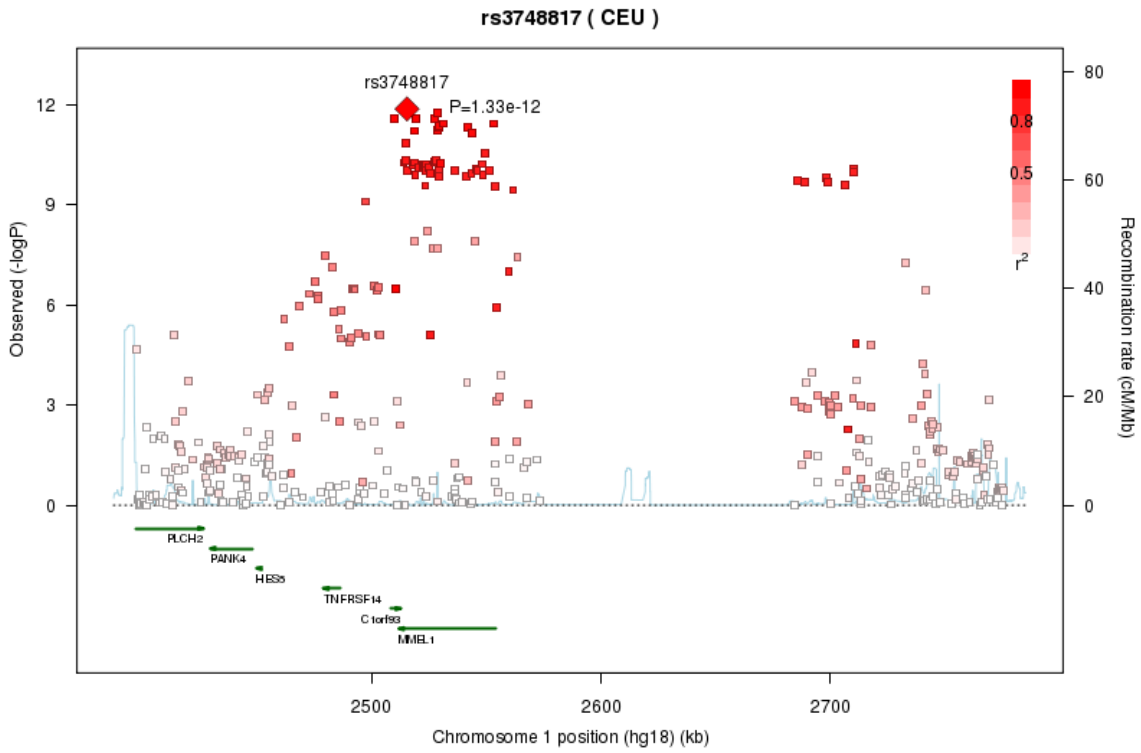
Stratum	OR	95% CI
AUSNZ	1.19	[0.93–1.53]
Belgium	1.00	[0.80–1.25]
Finland	1.05	[0.78–1.41]
France	0.83	[0.63–1.10]
Germany	1.12	[1.02–1.23]
Italy	1.22	[1.04–1.43]
Norway	1.08	[0.89–1.31]
Sweden	1.10	[0.98–1.23]
UK	1.16	[1.06–1.26]
US	1.08	[0.98–1.20]
Summary	**1.11**	**[1.07–1.16]**



A) Regional Association and B) Forest Plot

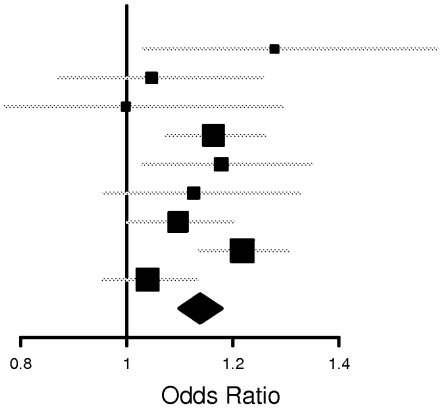
Supplementary Figure 48. Discovery phase rs3748817.

A



B

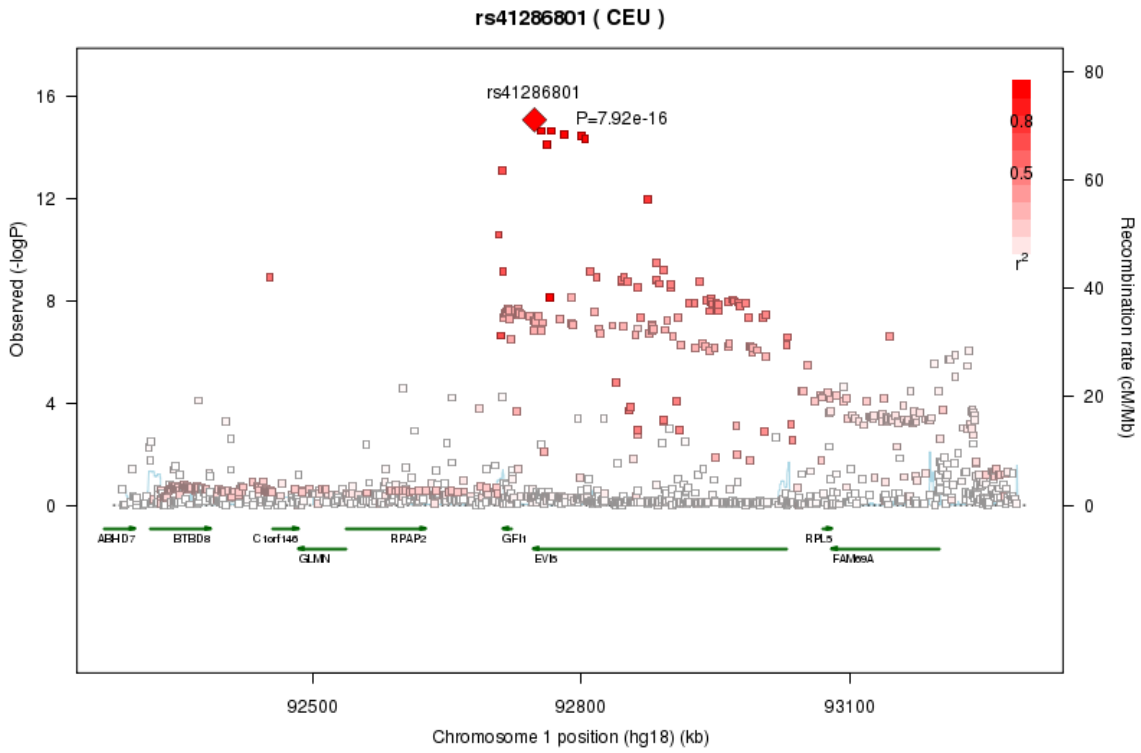
Stratum	OR	95% CI
AUSNZ	1.28	[1.03–1.58]
Belgium	1.05	[0.87–1.26]
Finland	1.00	[0.77–1.29]
Germany	1.16	[1.07–1.26]
Italy	1.18	[1.03–1.35]
Norway	1.13	[0.96–1.33]
Sweden	1.10	[1.00–1.20]
UK	1.22	[1.14–1.31]
US	1.04	[0.95–1.13]
Summary	**1.14**	**[1.10–1.18]**



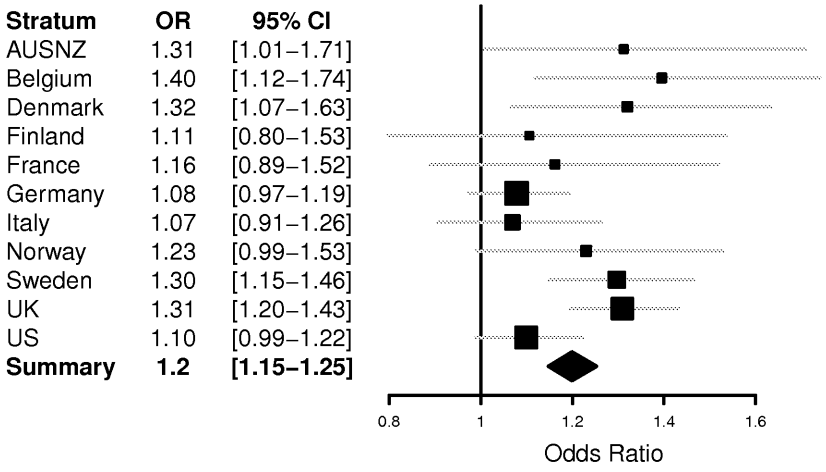
A) Regional Association and B) Forest Plot

Supplementary Figure 49. Discovery phase rs41286801.

A



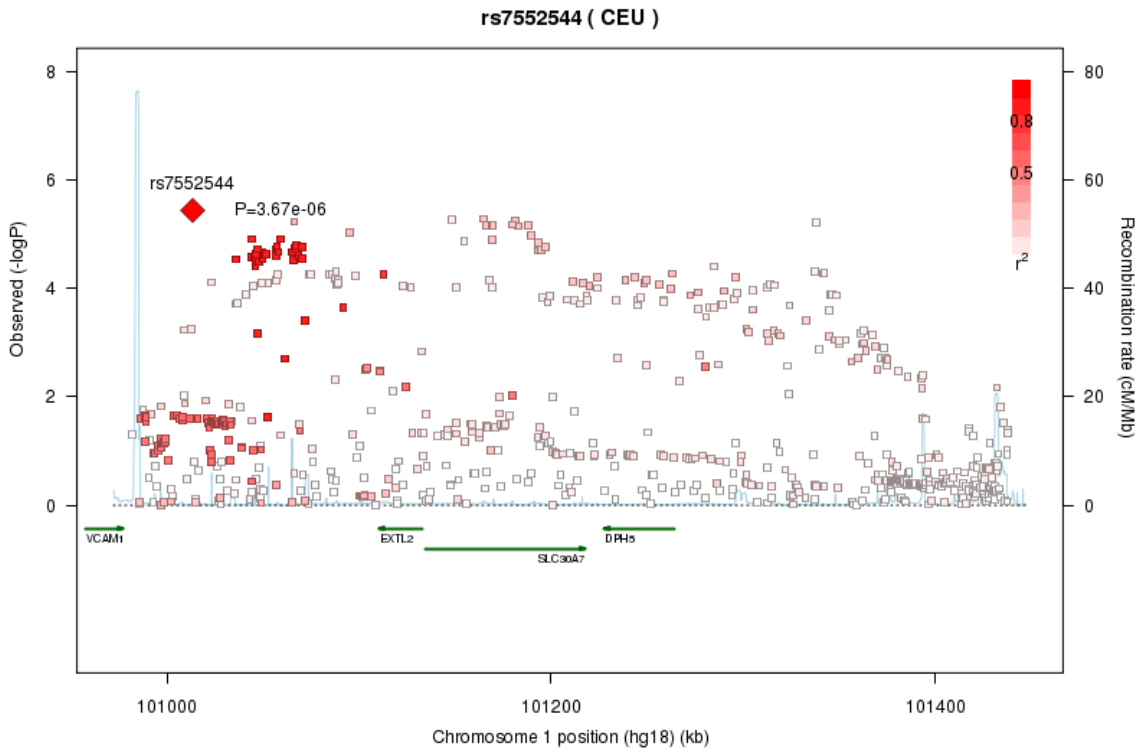
B



A) Regional Association and B) Forest Plot

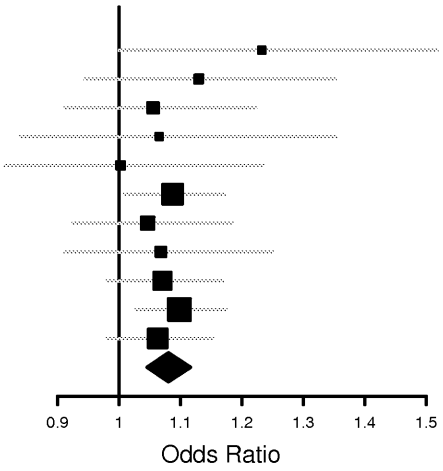
Supplementary Figure 50. Discovery phase rs7552544.

A



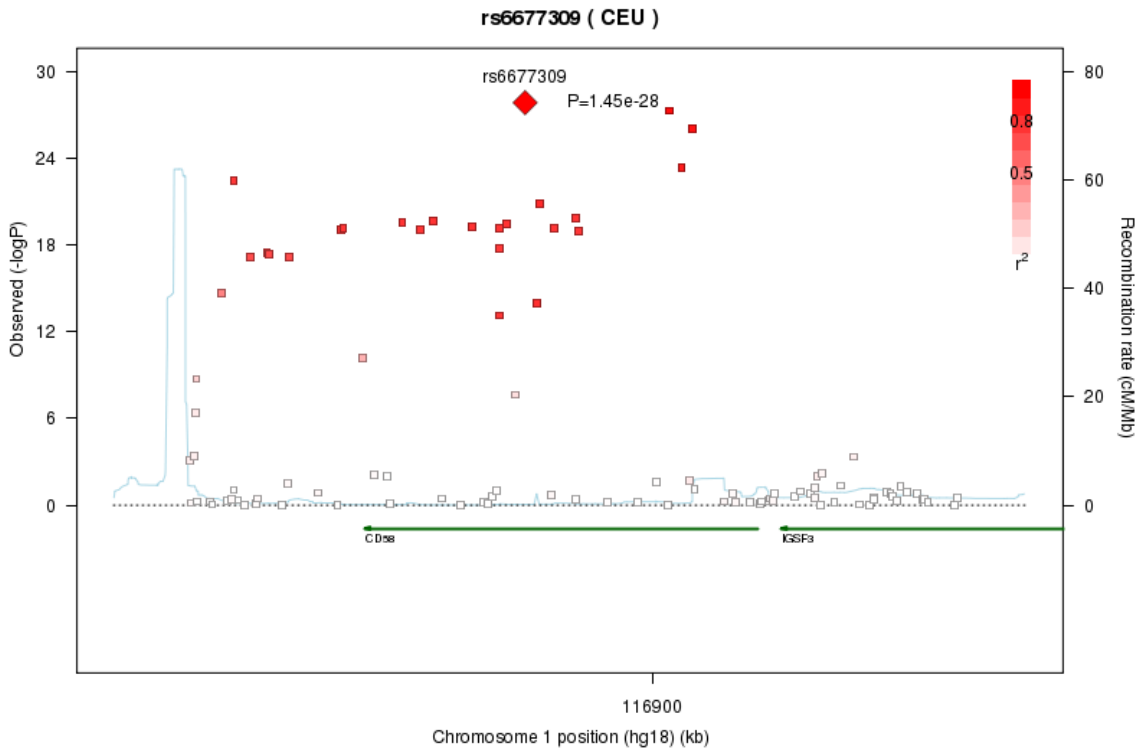
B

Stratum	OR	95% CI
AUSNZ	1.23	[1.00–1.52]
Belgium	1.13	[0.94–1.35]
Denmark	1.06	[0.91–1.22]
Finland	1.07	[0.84–1.35]
France	1.00	[0.81–1.23]
Germany	1.09	[1.01–1.17]
Italy	1.05	[0.93–1.18]
Norway	1.07	[0.91–1.25]
Sweden	1.07	[0.98–1.17]
UK	1.10	[1.03–1.17]
US	1.06	[0.98–1.15]
Summary	**1.08**	**[1.05–1.12]**

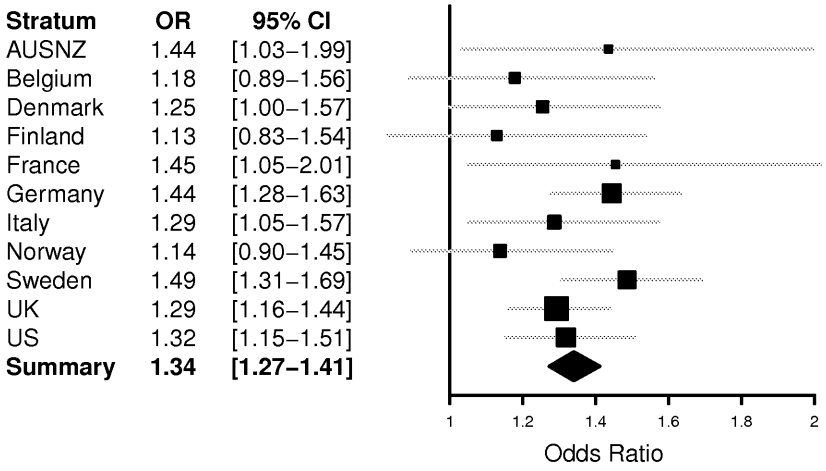


A) Regional Association and B) Forest Plot

Supplementary Figure 51. Discovery phase rs6677309.
A



B

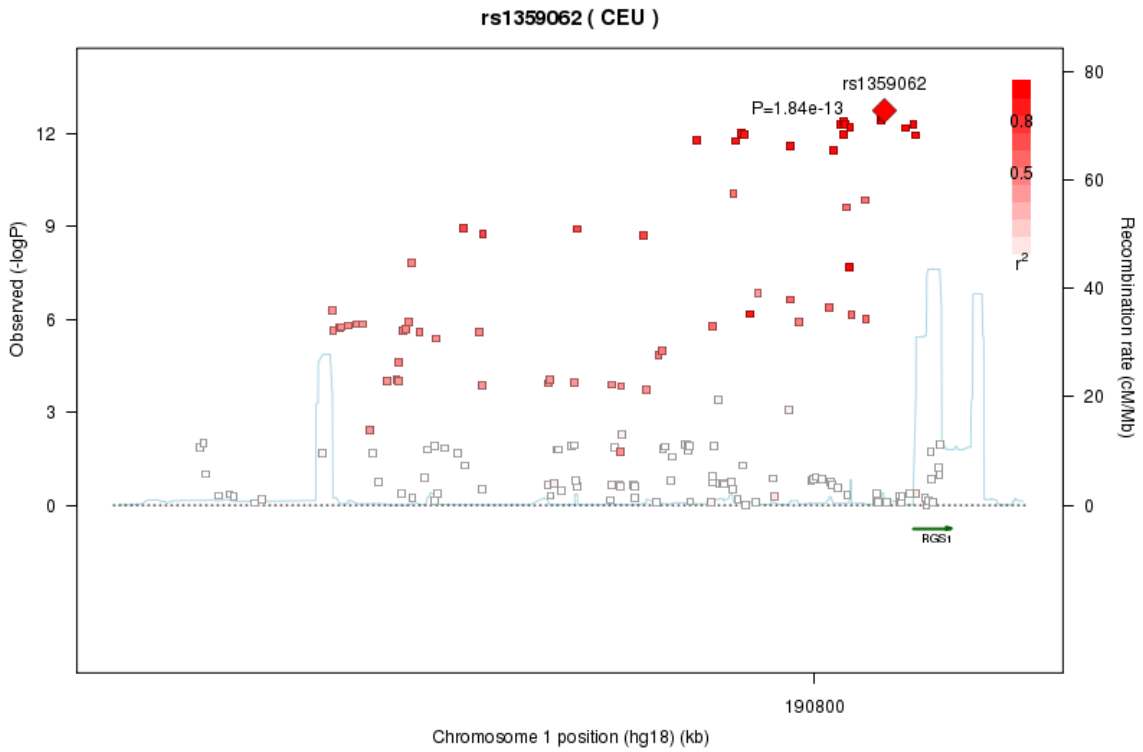


Stratum	OR	95% CI
AUSNZ	1.44	[1.03–1.99]
Belgium	1.18	[0.89–1.56]
Denmark	1.25	[1.00–1.57]
Finland	1.13	[0.83–1.54]
France	1.45	[1.05–2.01]
Germany	1.44	[1.28–1.63]
Italy	1.29	[1.05–1.57]
Norway	1.14	[0.90–1.45]
Sweden	1.49	[1.31–1.69]
UK	1.29	[1.16–1.44]
US	1.32	[1.15–1.51]
Summary	**1.34**	**[1.27–1.41]**

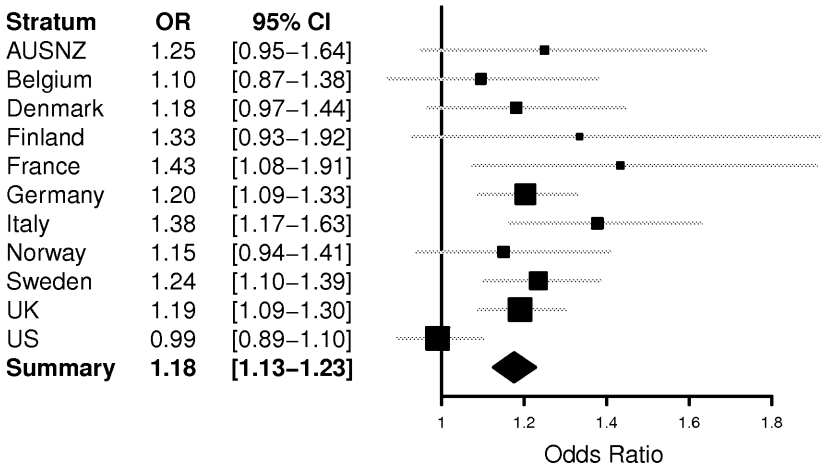
A) Regional Association and B) Forest Plot

Supplementary Figure 52. Discovery phase rs1359062.

A



B

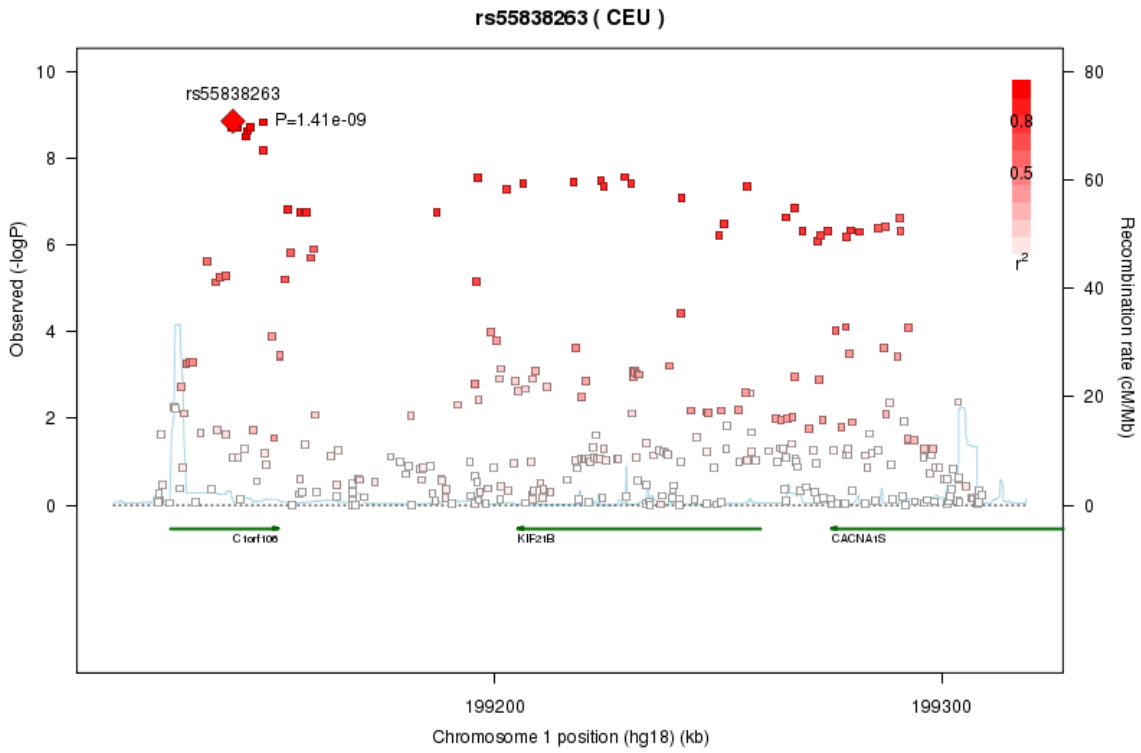


Stratum	OR	95% CI
AUSNZ	1.25	[0.95–1.64]
Belgium	1.10	[0.87–1.38]
Denmark	1.18	[0.97–1.44]
Finland	1.33	[0.93–1.92]
France	1.43	[1.08–1.91]
Germany	1.20	[1.09–1.33]
Italy	1.38	[1.17–1.63]
Norway	1.15	[0.94–1.41]
Sweden	1.24	[1.10–1.39]
UK	1.19	[1.09–1.30]
US	0.99	[0.89–1.10]
Summary	**1.18**	**[1.13–1.23]**

A) Regional Association and B) Forest Plot

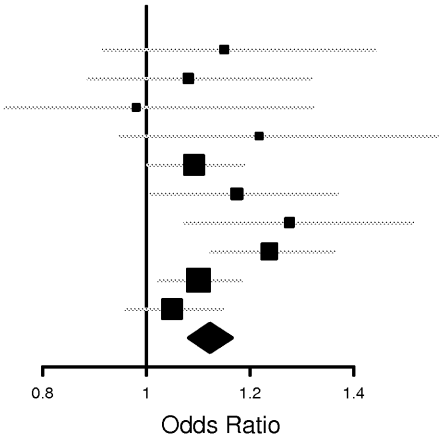
Supplementary Figure 53: Discovery phase rs55838263.

A



B

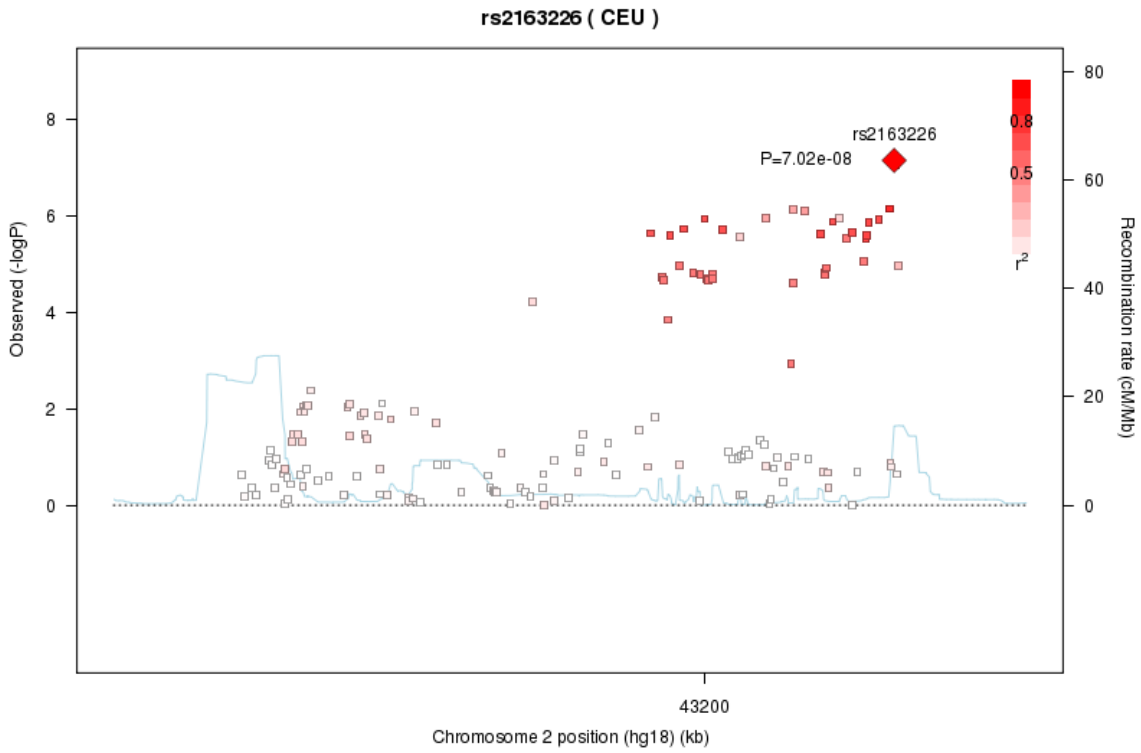
Stratum	OR	95% CI
AUSNZ	1.15	[0.92–1.44]
Belgium	1.08	[0.89–1.32]
Finland	0.98	[0.73–1.32]
France	1.22	[0.95–1.56]
Germany	1.09	[1.00–1.19]
Italy	1.17	[1.01–1.37]
Norway	1.28	[1.07–1.52]
Sweden	1.24	[1.12–1.36]
UK	1.10	[1.02–1.18]
US	1.05	[0.96–1.15]
Summary	**1.12**	**[1.08–1.17]**



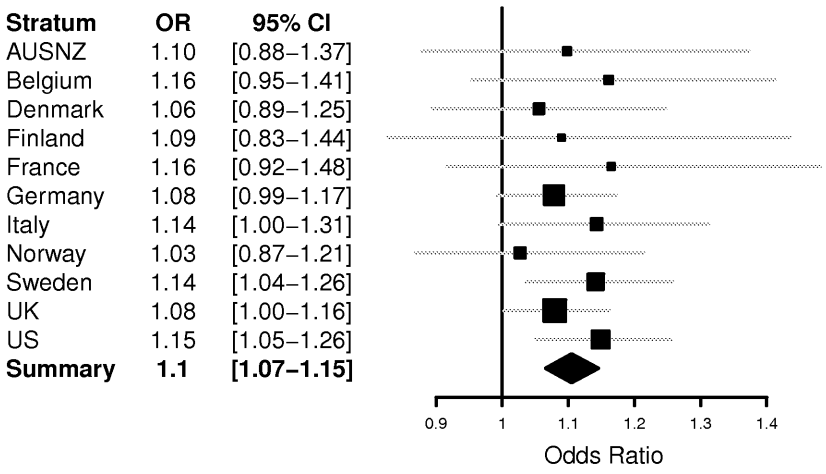
A) Regional Association and B) Forest Plot

Supplementary Figure 54. Discovery phase rs2163226.

A



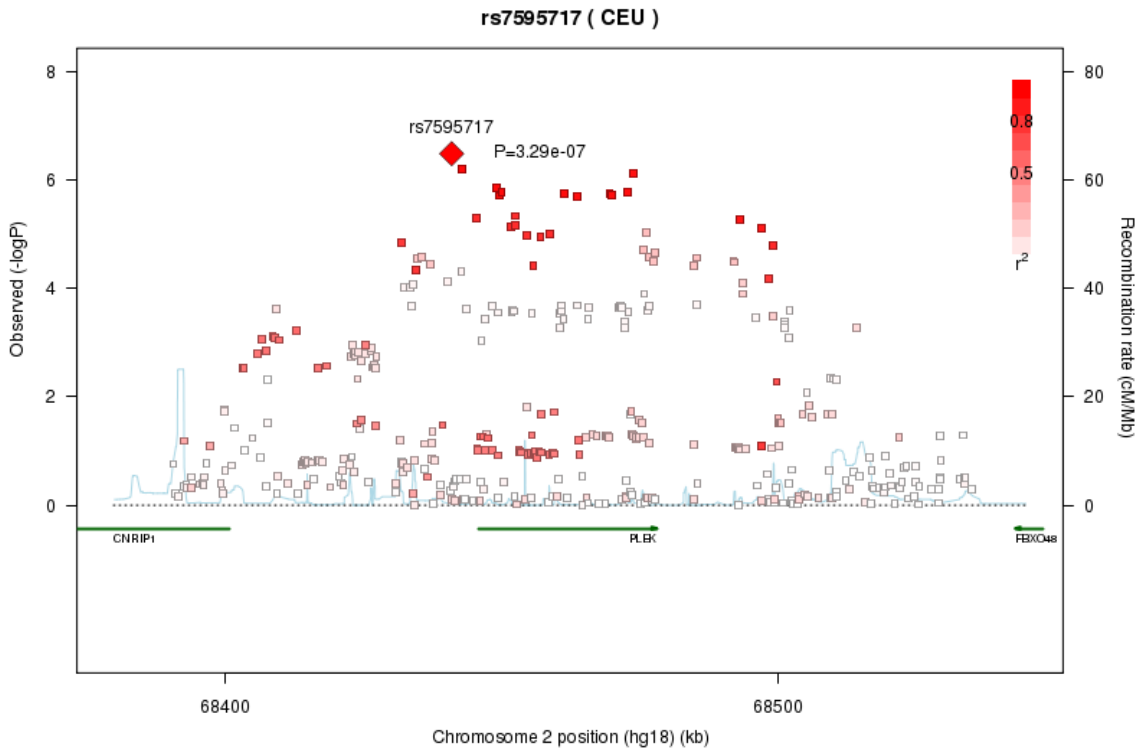
B



A) Regional Association and B) Forest Plot

Supplementary Figure 55. Discovery phase rs7595717.

A



B

Stratum	OR	95% CI
AUSNZ	0.96	[0.76–1.20]
Belgium	1.02	[0.84–1.25]
Denmark	0.96	[0.81–1.13]
Finland	1.13	[0.87–1.46]
France	0.94	[0.74–1.20]
Germany	1.20	[1.11–1.31]
Italy	1.08	[0.94–1.24]
Norway	1.13	[0.96–1.33]
Sweden	1.13	[1.03–1.25]
UK	1.06	[0.99–1.14]
US	1.12	[1.02–1.22]
Summary	**1.1**	**[1.06–1.14]**

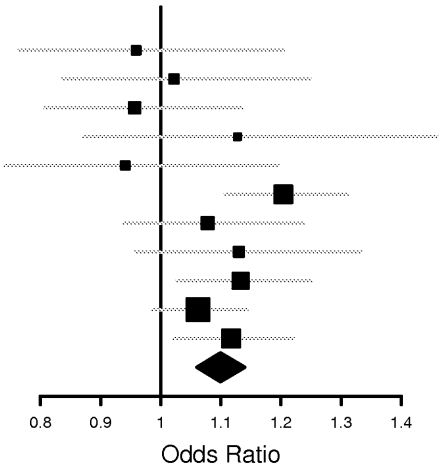
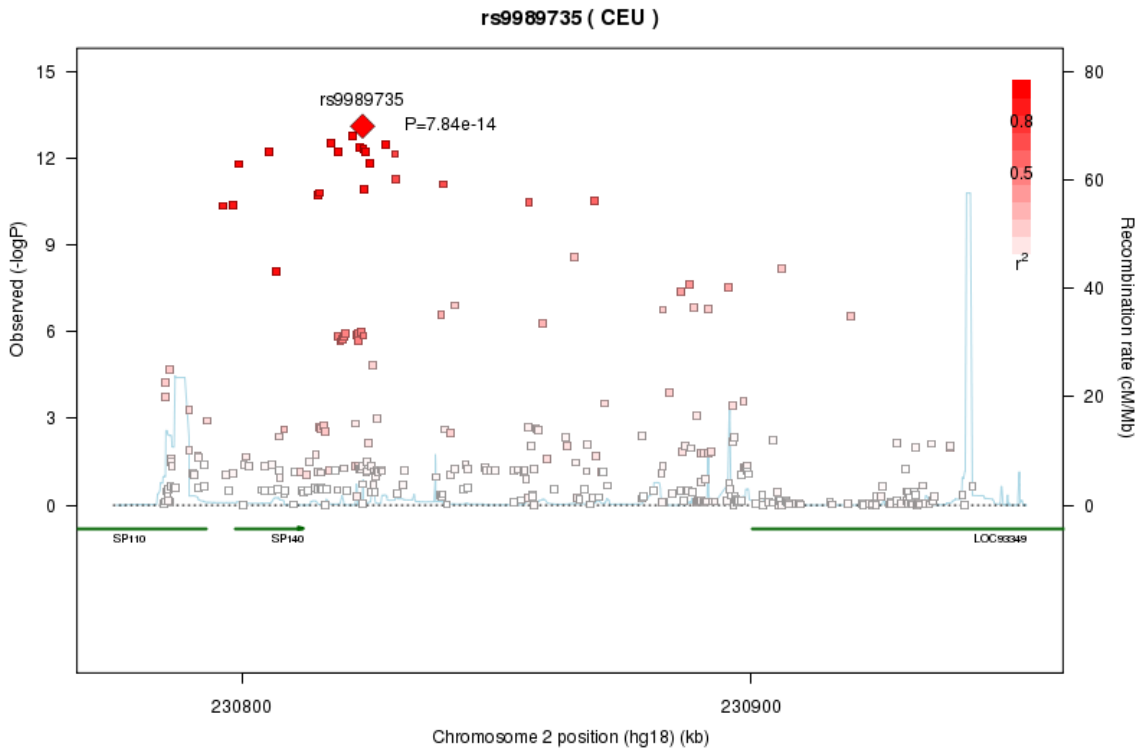


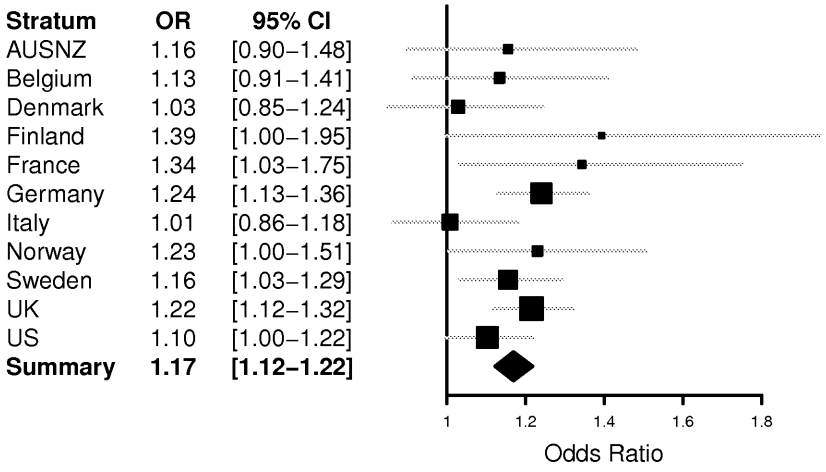
Figure S59. a) Regional Association and b) Forest Plot for rs7595717

Supplementary Figure 56. Discovery phase rs9989735.

A



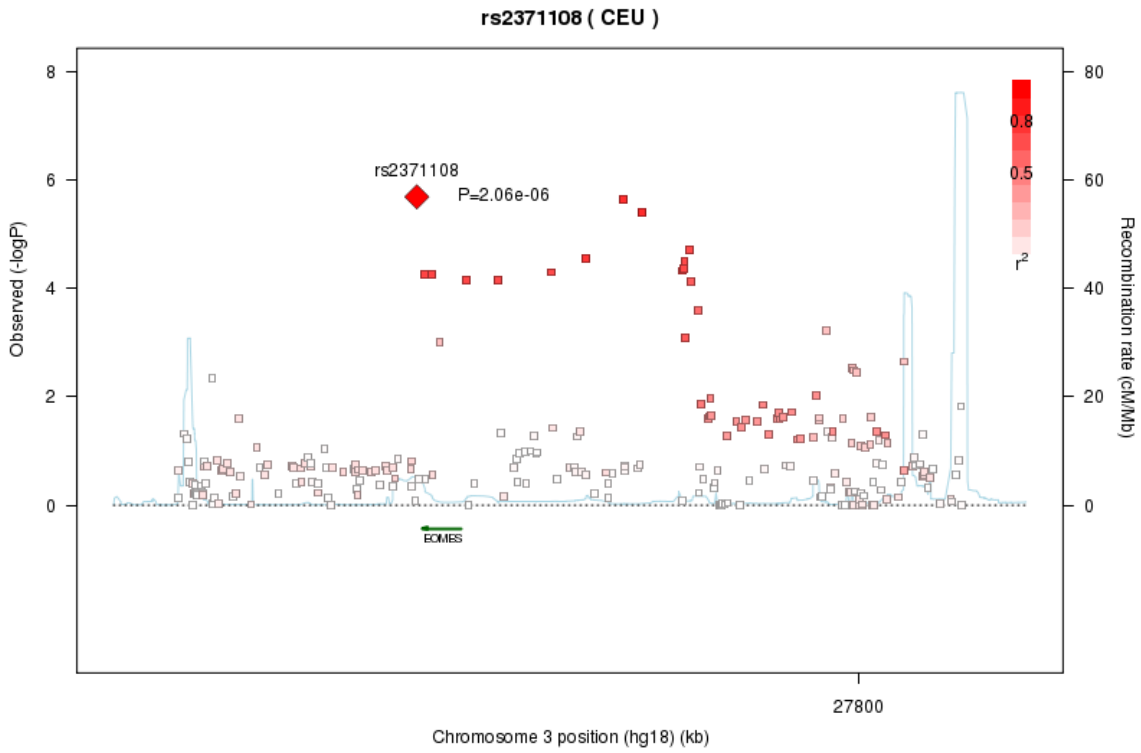
B



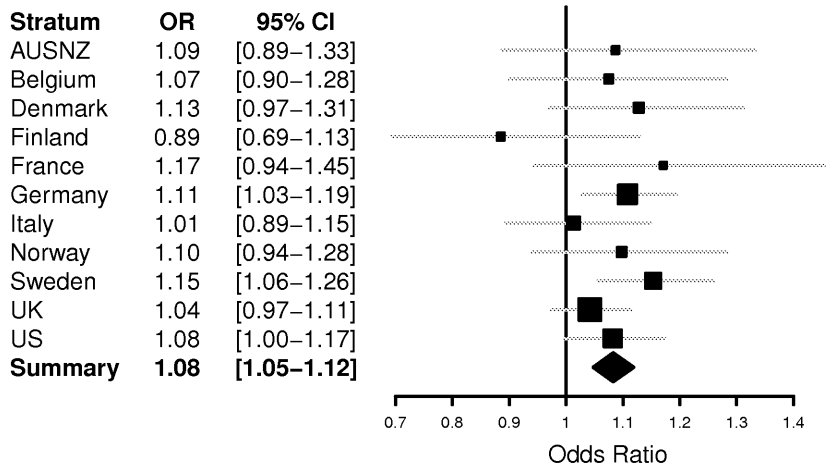
A) Regional Association and B) Forest Plot

Supplementary Figure 57. Discovery phase rs2371108.

A



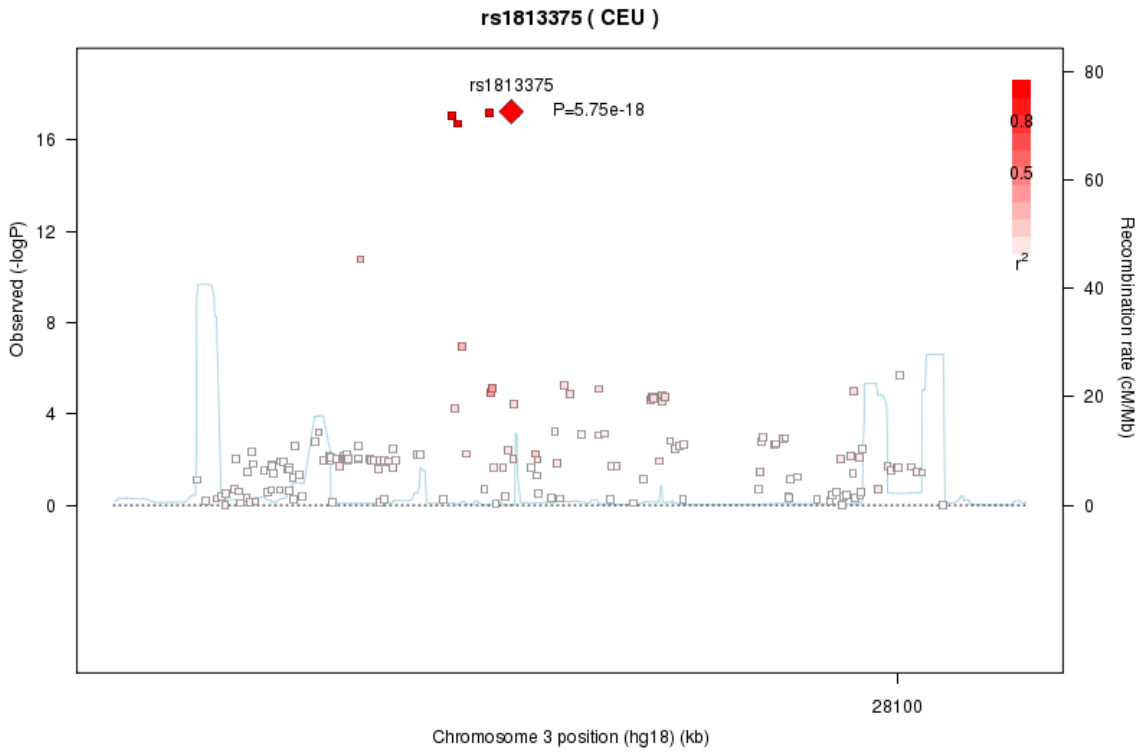
B



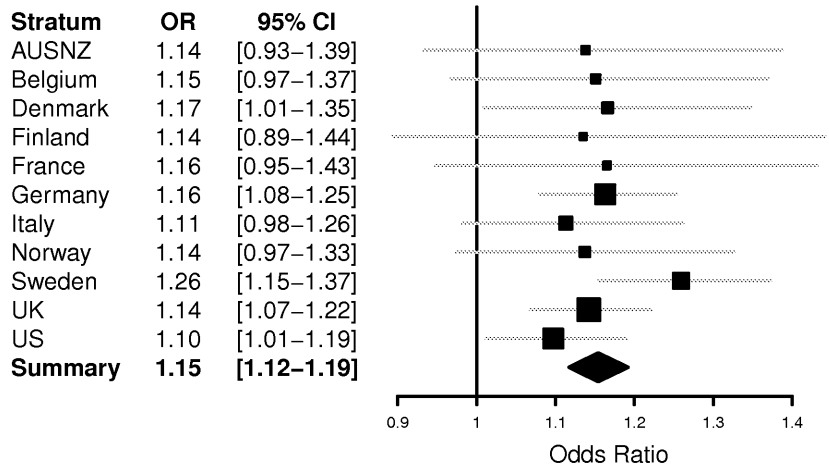
A) Regional Association and B) Forest Plot

Supplementary Figure 58. Discovery phase rs1813375.

A



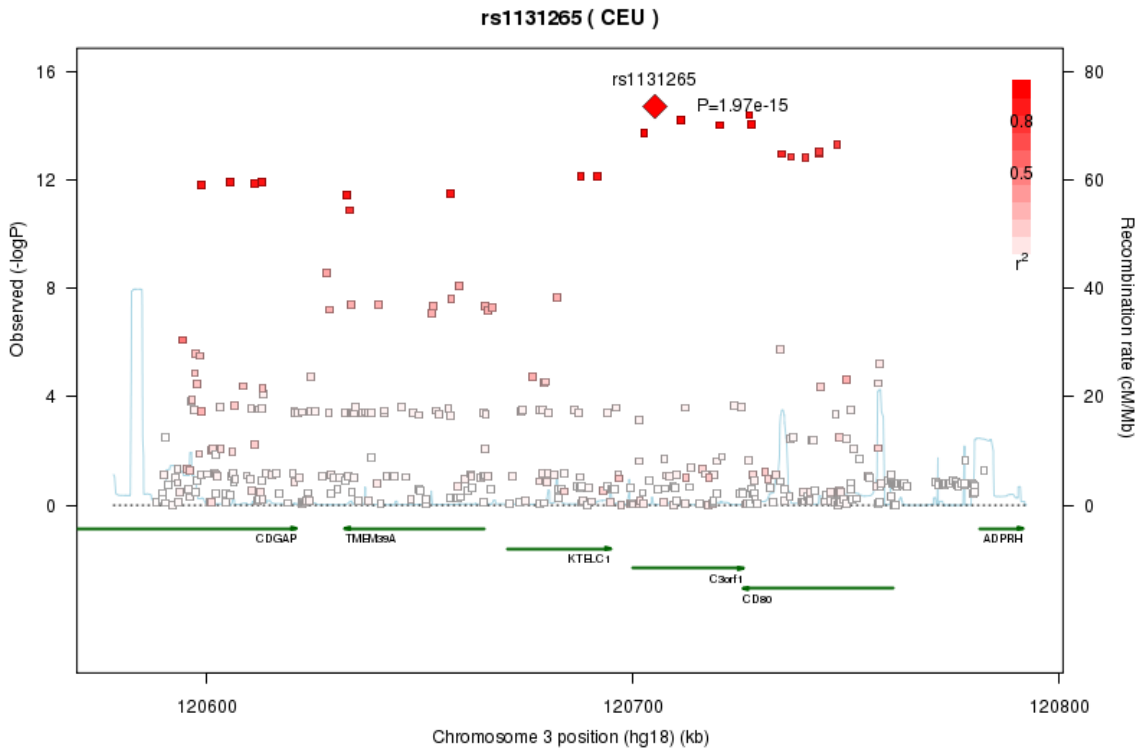
B



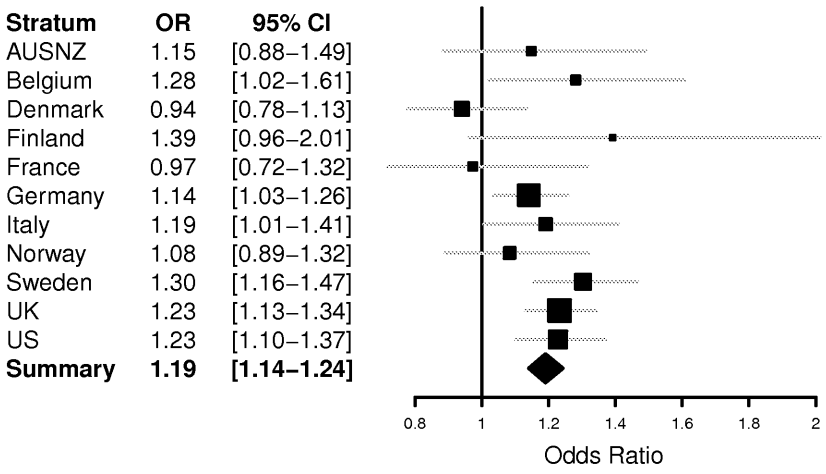
A) Regional Association and B) Forest Plot

Supplementary Figure 59. Discovery phase rs1131265.

A



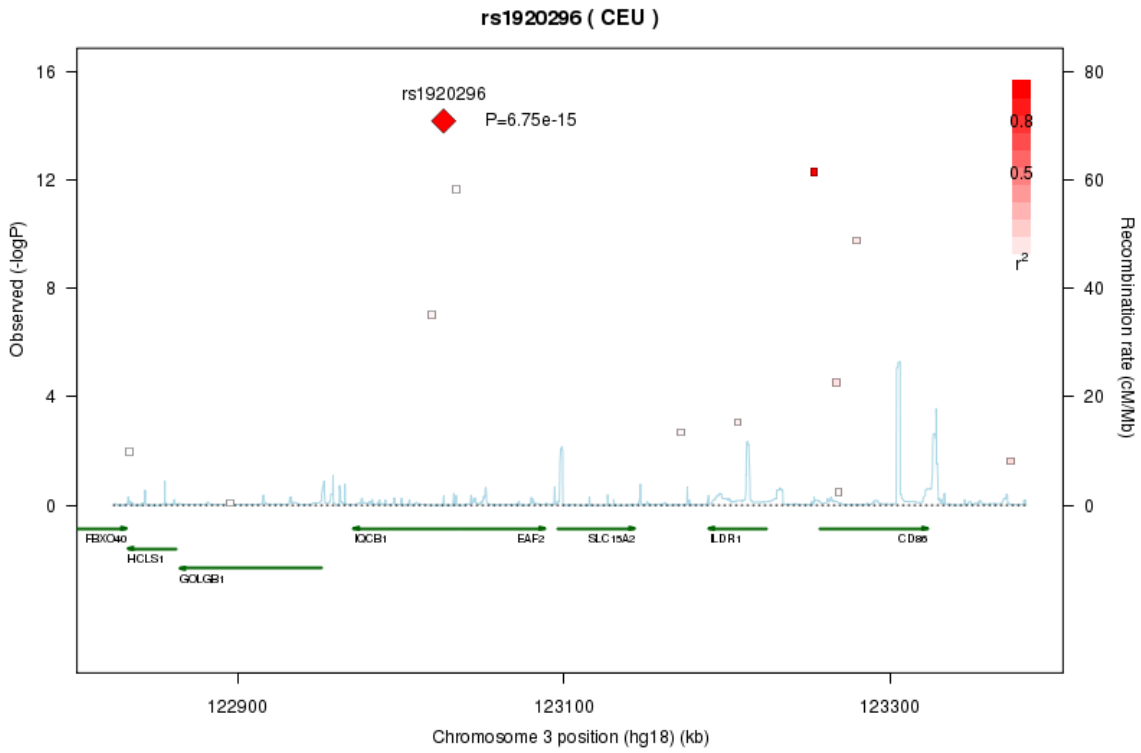
B



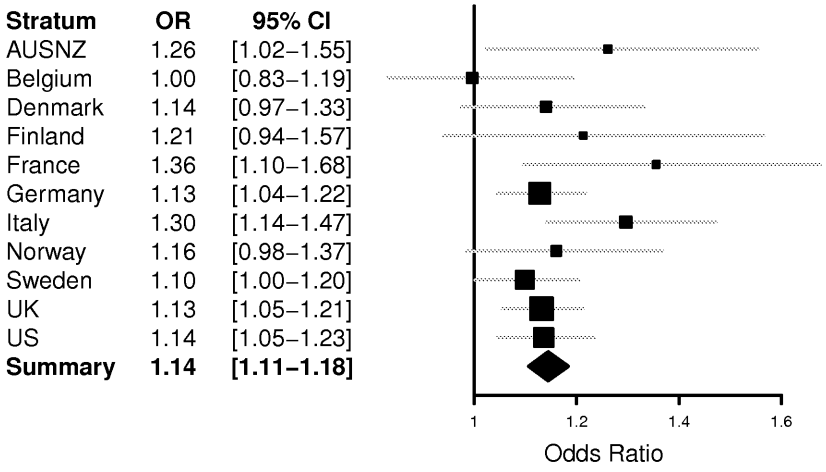
A) Regional Association and B) Forest Plot

Supplementary Figure 60. Discovery phase rs1920296.

A

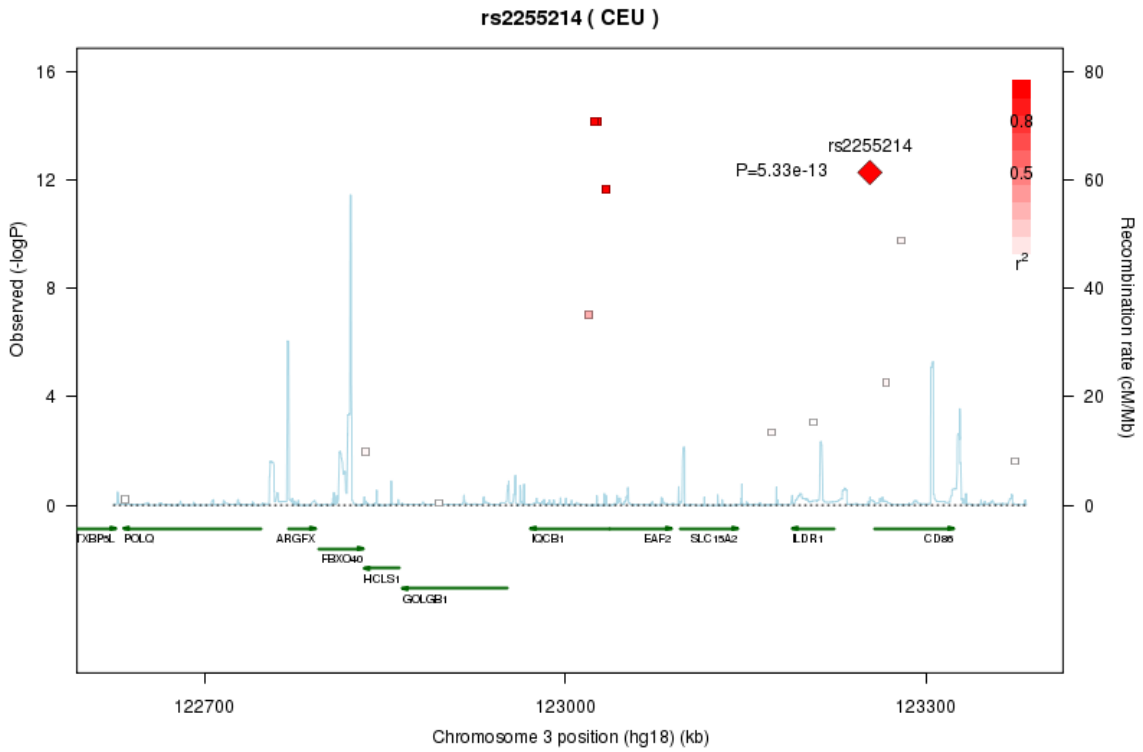


B

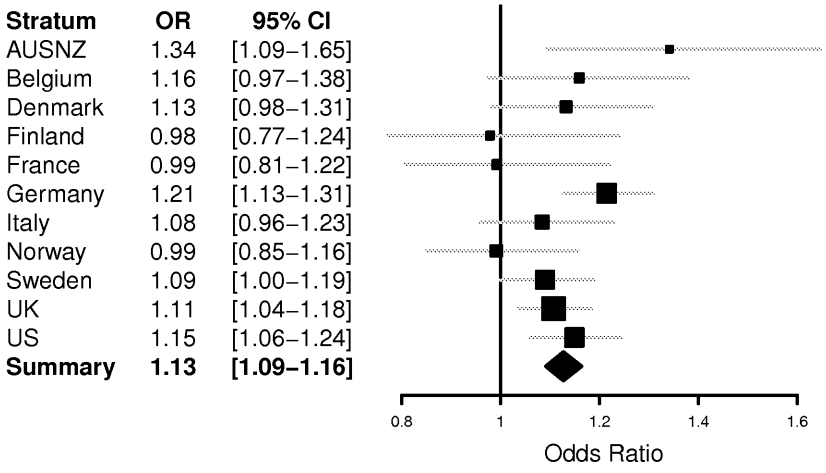


A) Regional Association and B) Forest Plot

Supplementary Figure 61. Discovery phase rs2255214.
A



B

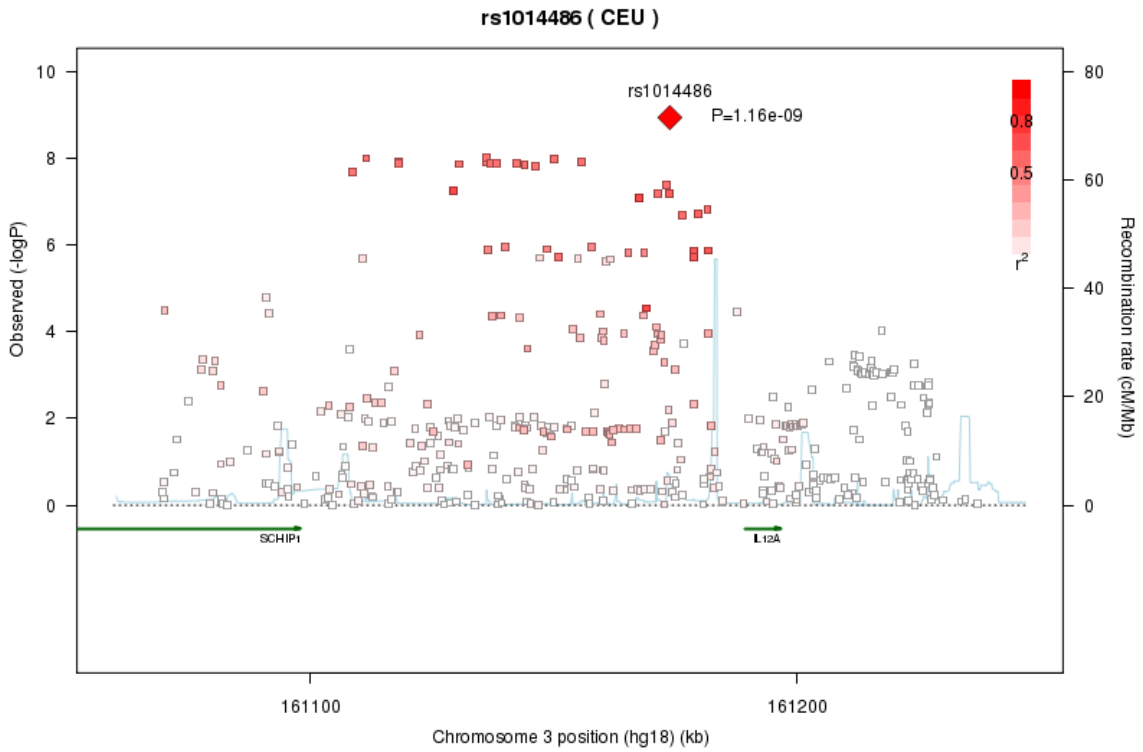


Stratum	OR	95% CI
AUSNZ	1.34	[1.09–1.65]
Belgium	1.16	[0.97–1.38]
Denmark	1.13	[0.98–1.31]
Finland	0.98	[0.77–1.24]
France	0.99	[0.81–1.22]
Germany	1.21	[1.13–1.31]
Italy	1.08	[0.96–1.23]
Norway	0.99	[0.85–1.16]
Sweden	1.09	[1.00–1.19]
UK	1.11	[1.04–1.18]
US	1.15	[1.06–1.24]
Summary	**1.13**	**[1.09–1.16]**

A) Regional Association and B) Forest Plot

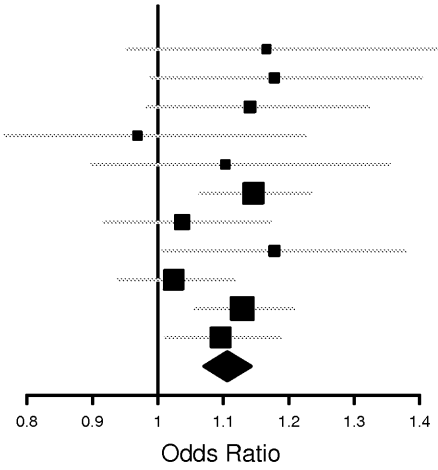
Supplementary Figure 62. Discovery phase rs1014486.

A



B

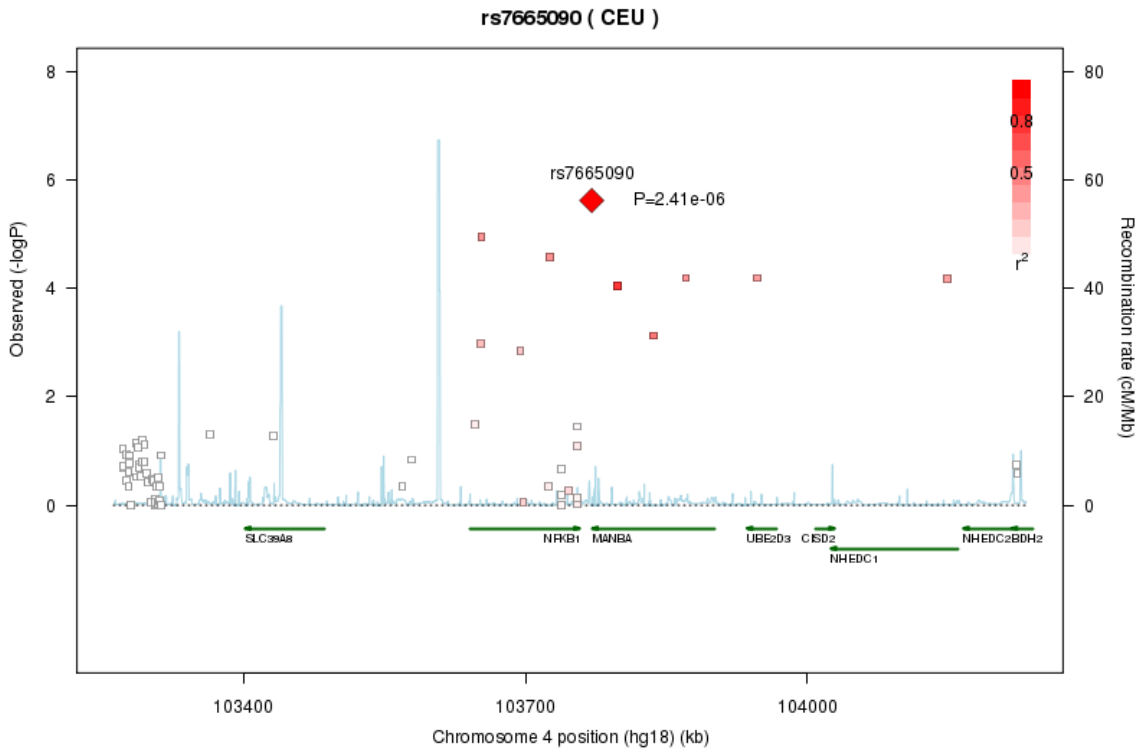
Stratum	OR	95% CI
AUSNZ	1.17	[0.95–1.43]
Belgium	1.18	[0.99–1.40]
Denmark	1.14	[0.98–1.32]
Finland	0.97	[0.77–1.22]
France	1.10	[0.90–1.35]
Germany	1.15	[1.06–1.23]
Italy	1.04	[0.92–1.17]
Norway	1.18	[1.01–1.38]
Sweden	1.02	[0.94–1.12]
UK	1.13	[1.06–1.21]
US	1.10	[1.01–1.19]
Summary	**1.11**	**[1.07–1.14]**



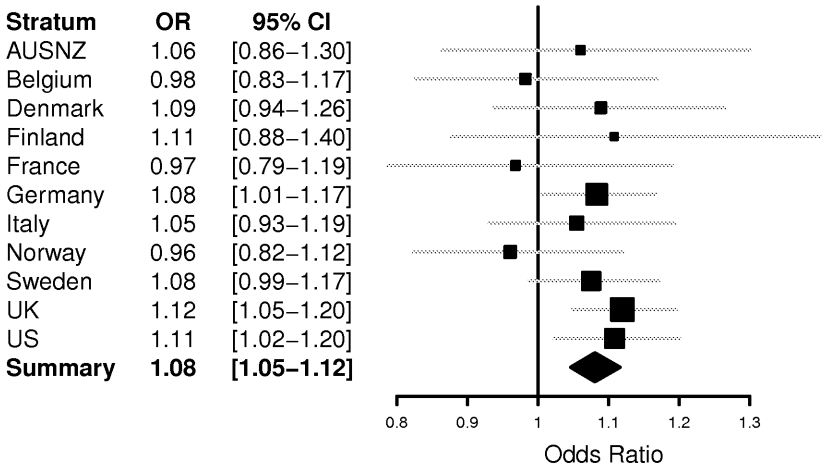
A) Regional Association and B) Forest Plot

Supplementary Figure 63. Discovery phase rs7665090.

A



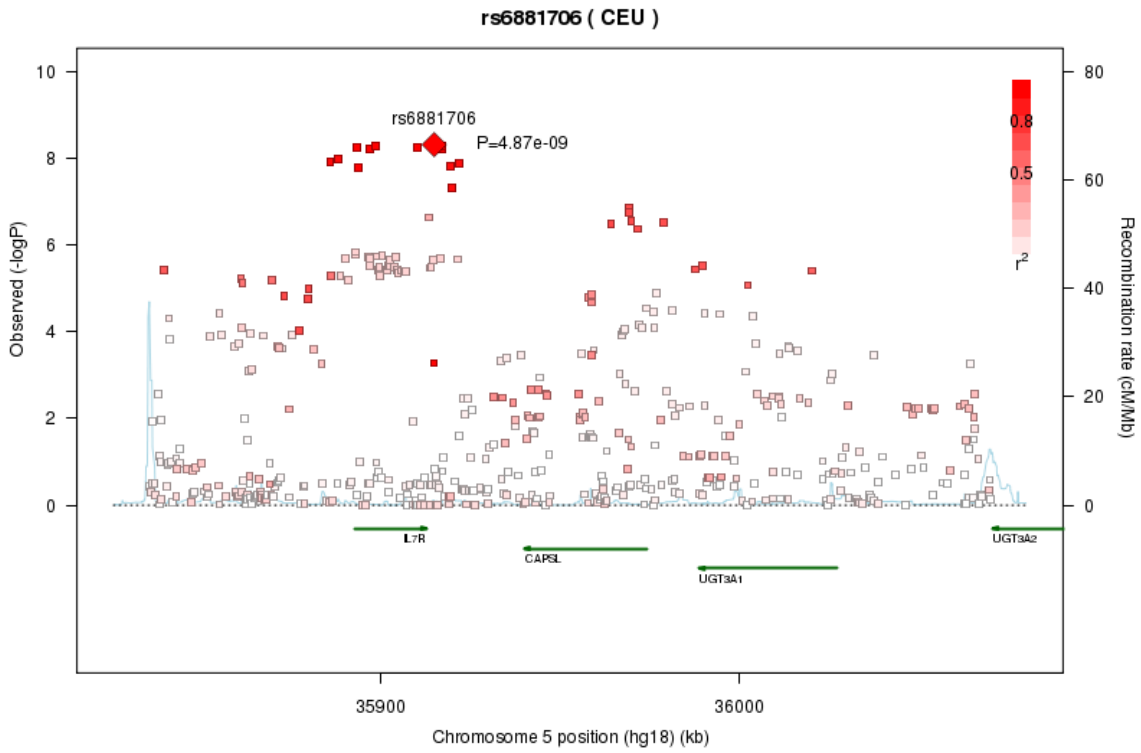
B



A) Regional Association and B) Forest Plot

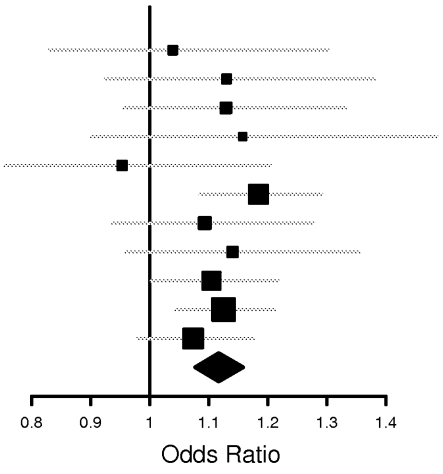
Supplementary Figure 64. Discovery phase rs6881706.

A



B

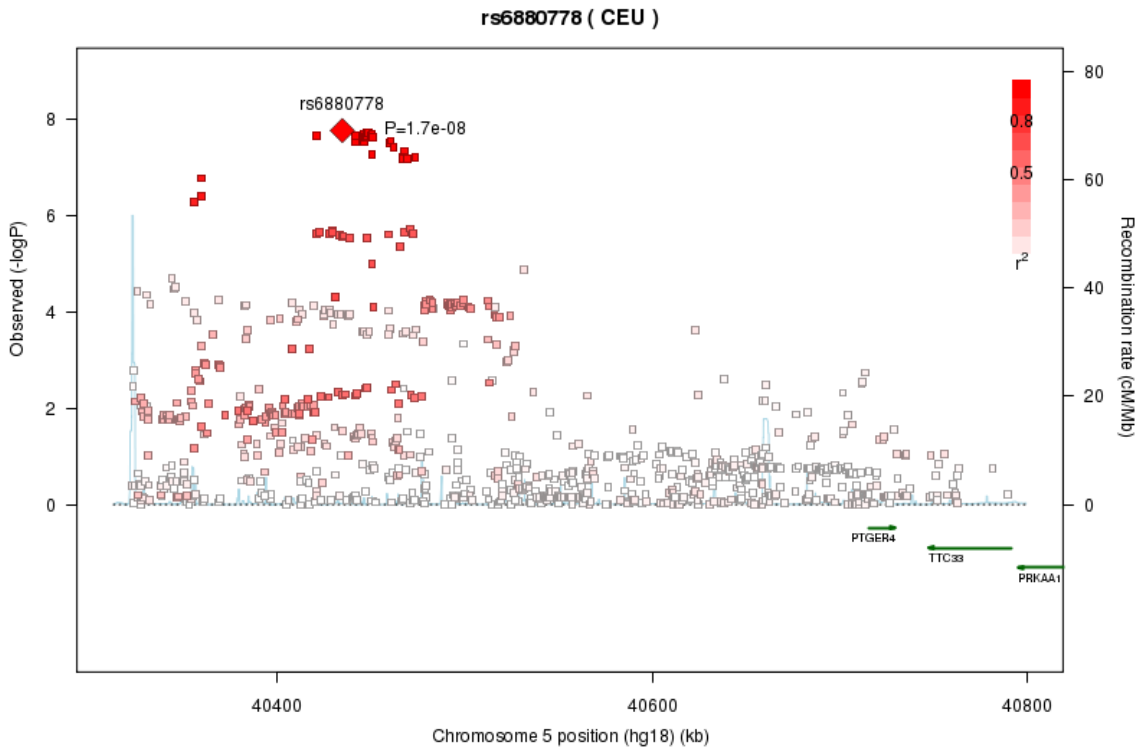
Stratum	OR	95% CI
AUSNZ	1.04	[0.83–1.30]
Belgium	1.13	[0.93–1.38]
Denmark	1.13	[0.96–1.33]
Finland	1.16	[0.90–1.49]
France	0.95	[0.75–1.20]
Germany	1.18	[1.09–1.29]
Italy	1.09	[0.94–1.28]
Norway	1.14	[0.96–1.36]
Sweden	1.10	[1.00–1.22]
UK	1.13	[1.04–1.21]
US	1.07	[0.98–1.18]
Summary	**1.12**	**[1.08–1.16]**



A) Regional Association and B) Forest Plot for rs6881706

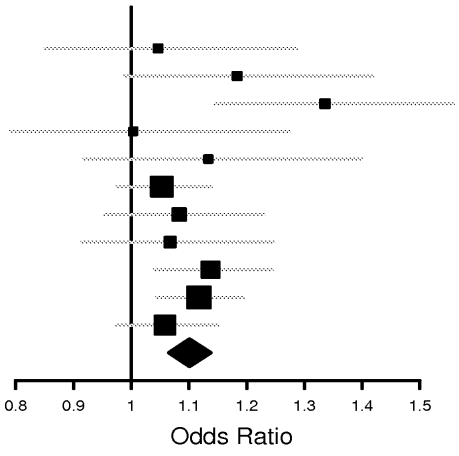
Supplementary Figure 65. Discovery phase rs6880778.

A



B

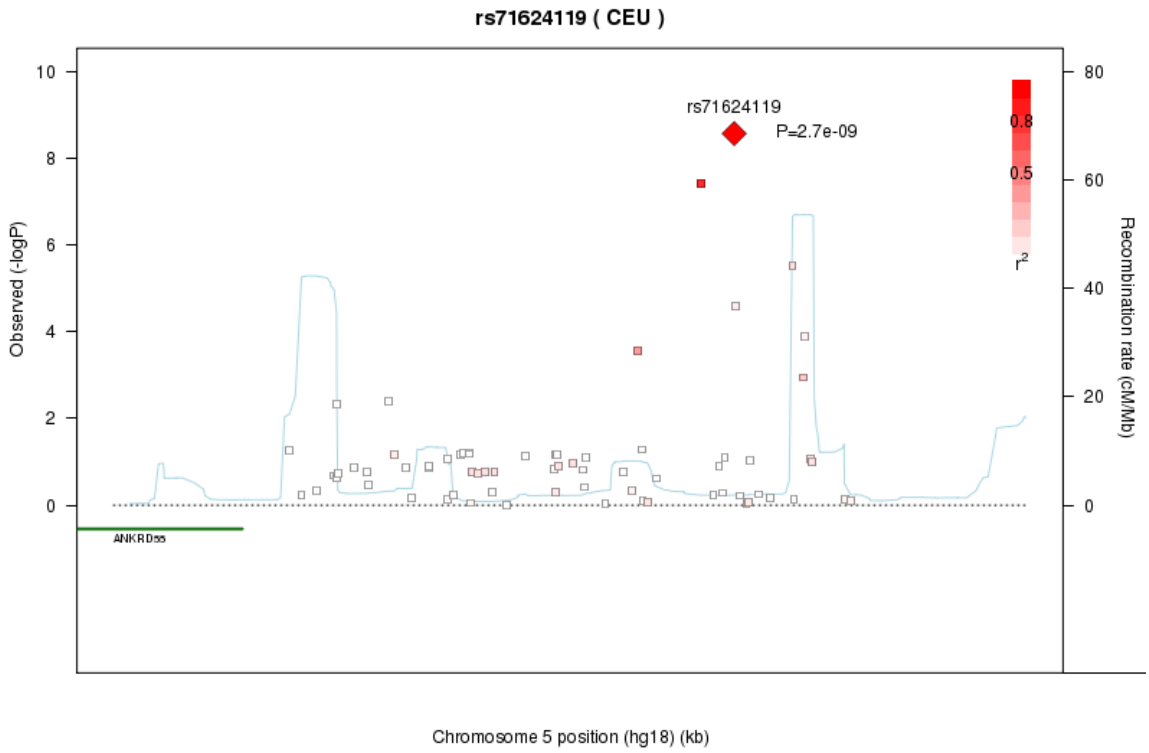
Stratum	OR	95% CI
AUSNZ	1.05	[0.85–1.29]
Belgium	1.18	[0.99–1.42]
Denmark	1.34	[1.14–1.56]
Finland	1.00	[0.79–1.27]
France	1.13	[0.92–1.40]
Germany	1.05	[0.97–1.14]
Italy	1.08	[0.95–1.23]
Norway	1.07	[0.91–1.25]
Sweden	1.14	[1.04–1.24]
UK	1.12	[1.04–1.20]
US	1.06	[0.97–1.15]
Summary	**1.10**	**[1.06–1.14]**



A) Regional Association and B) Forest Plot

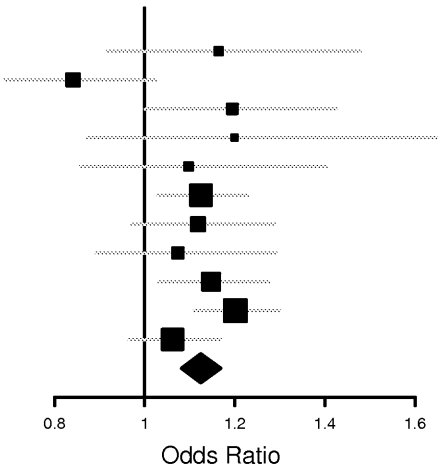
Supplementary Figure 66. Discovery phase rs71624119.

A



B

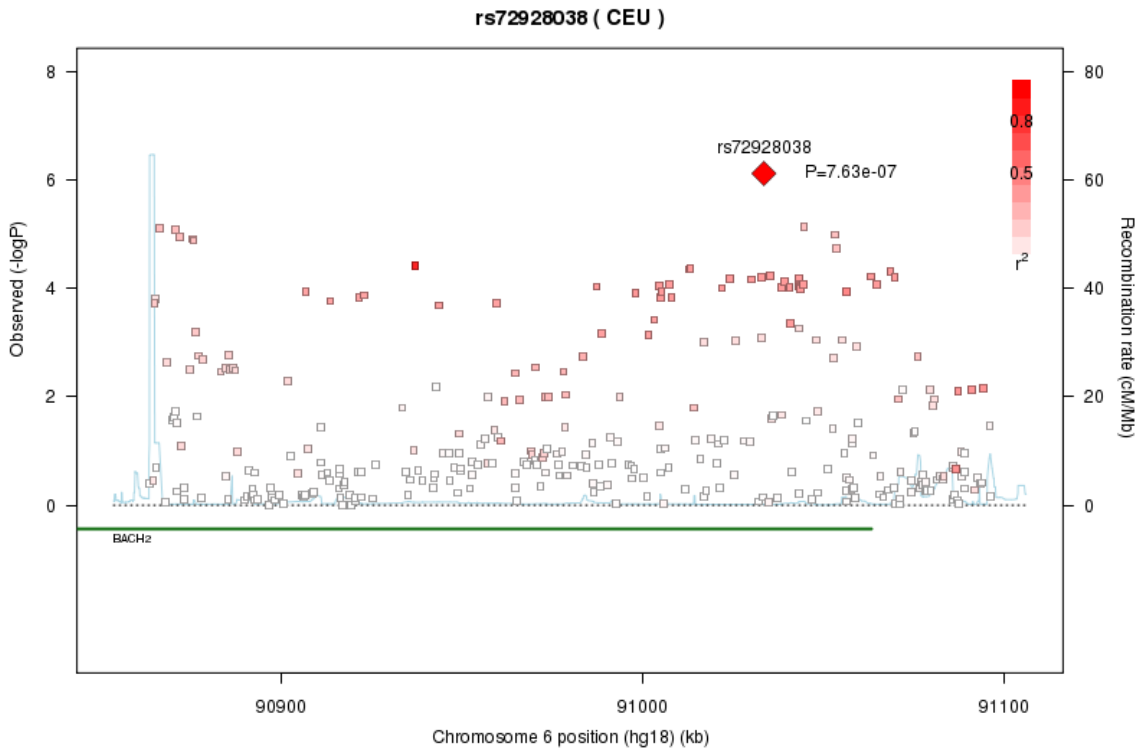
Stratum	OR	95% CI
AUSNZ	1.16	[0.92–1.48]
Belgium	0.84	[0.69–1.03]
Denmark	1.20	[1.00–1.43]
Finland	1.20	[0.87–1.65]
France	1.10	[0.86–1.40]
Germany	1.13	[1.03–1.23]
Italy	1.12	[0.97–1.29]
Norway	1.07	[0.89–1.29]
Sweden	1.15	[1.03–1.28]
UK	1.20	[1.11–1.30]
US	1.06	[0.97–1.17]
Summary	**1.12**	**[1.08–1.17]**



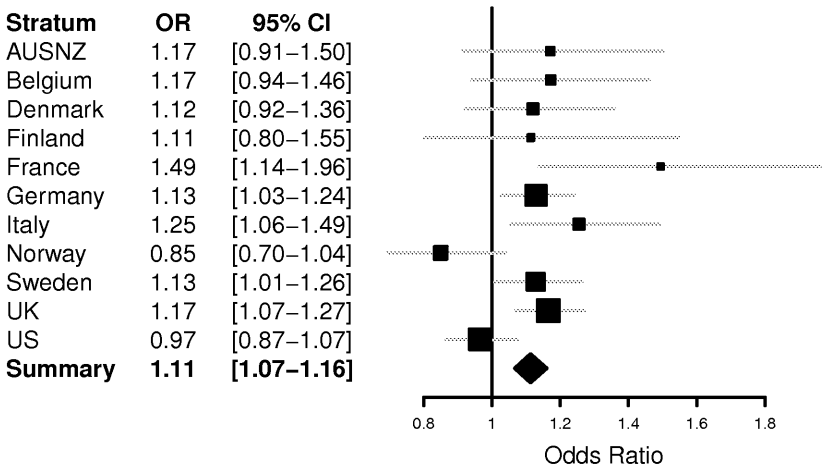
A) Regional Association and B) Forest Plot

Supplementary Figure 67. Discovery phase rs72928038.

A



B

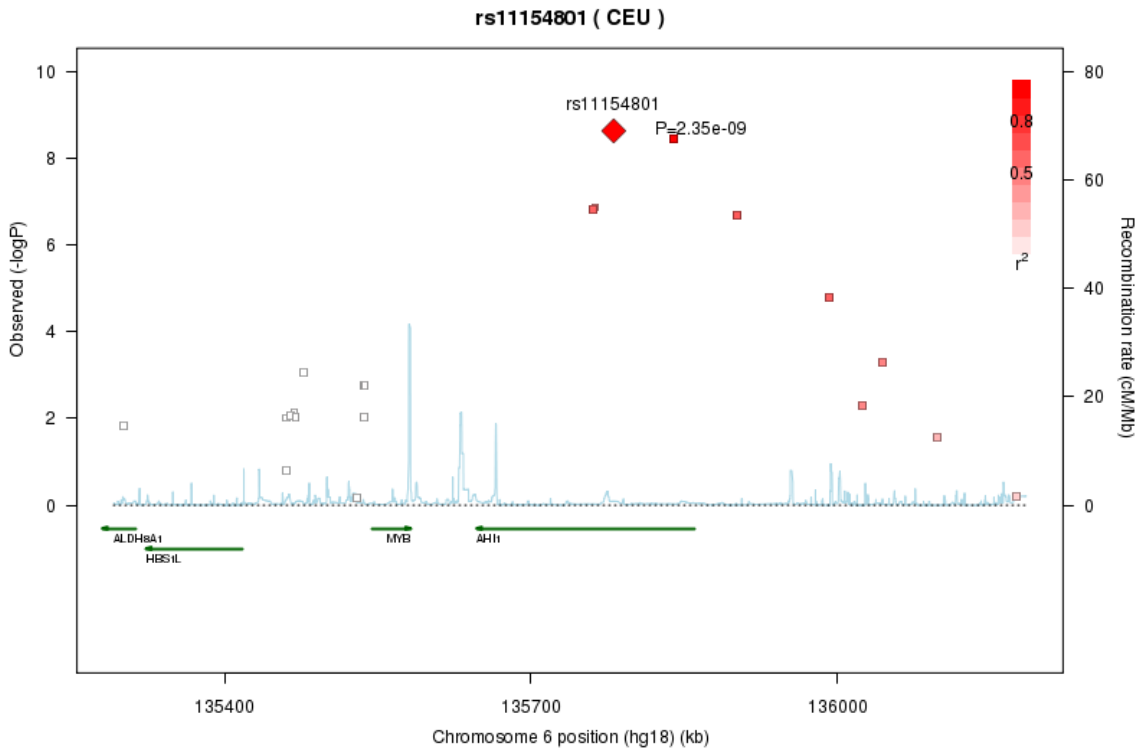


Stratum	OR	95% CI
AUSNZ	1.17	[0.91–1.50]
Belgium	1.17	[0.94–1.46]
Denmark	1.12	[0.92–1.36]
Finland	1.11	[0.80–1.55]
France	1.49	[1.14–1.96]
Germany	1.13	[1.03–1.24]
Italy	1.25	[1.06–1.49]
Norway	0.85	[0.70–1.04]
Sweden	1.13	[1.01–1.26]
UK	1.17	[1.07–1.27]
US	0.97	[0.87–1.07]
Summary	**1.11**	**[1.07–1.16]**

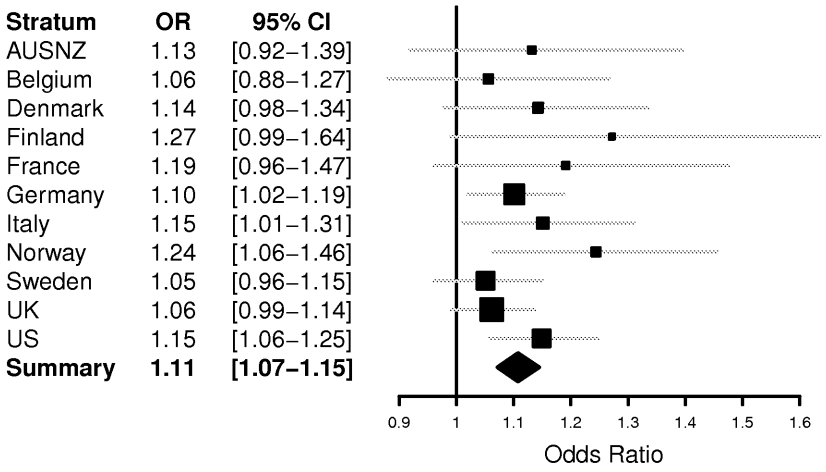
A) Regional Association and B) Forest Plot

Supplementary Figure 68. Discovery phase rs11154801.

A



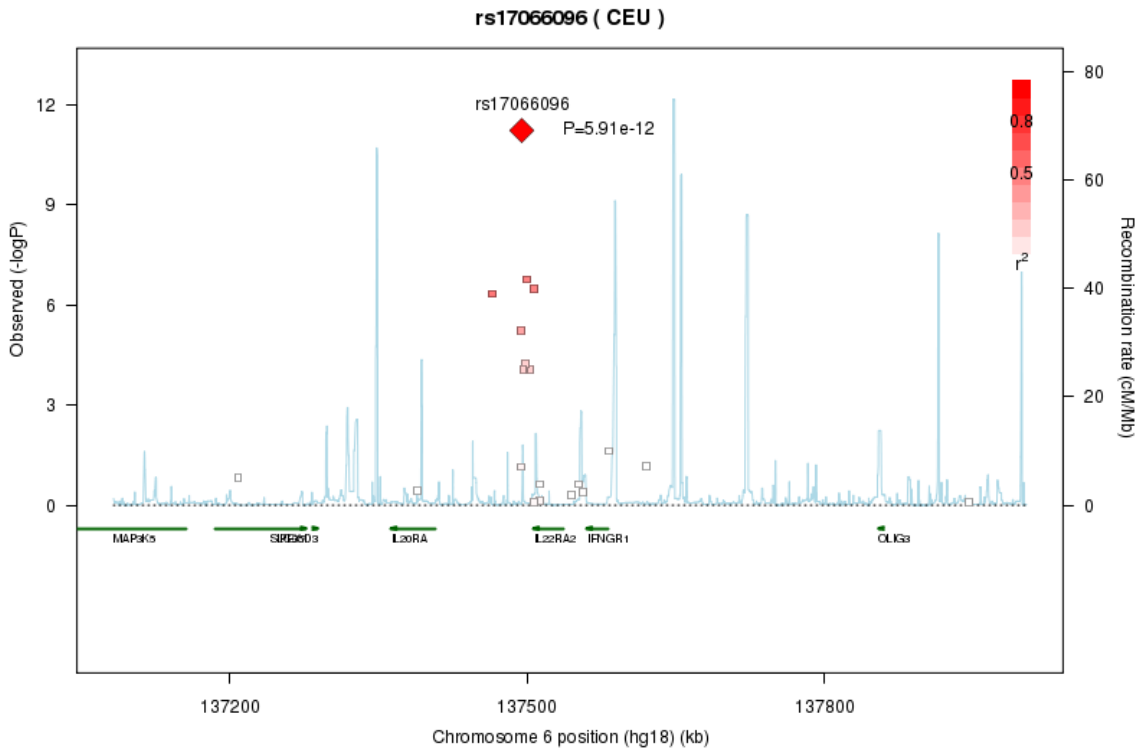
B



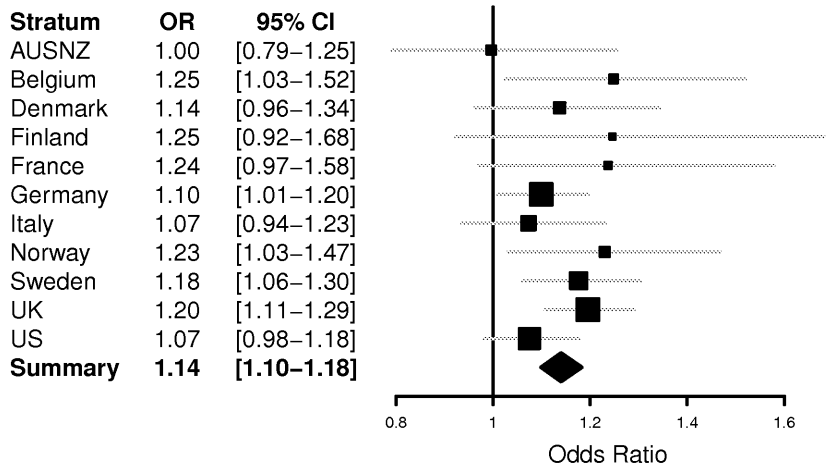
A) Regional Association and B) Forest Plot

Supplementary Figure 69. Discovery phase rs17066096.

A



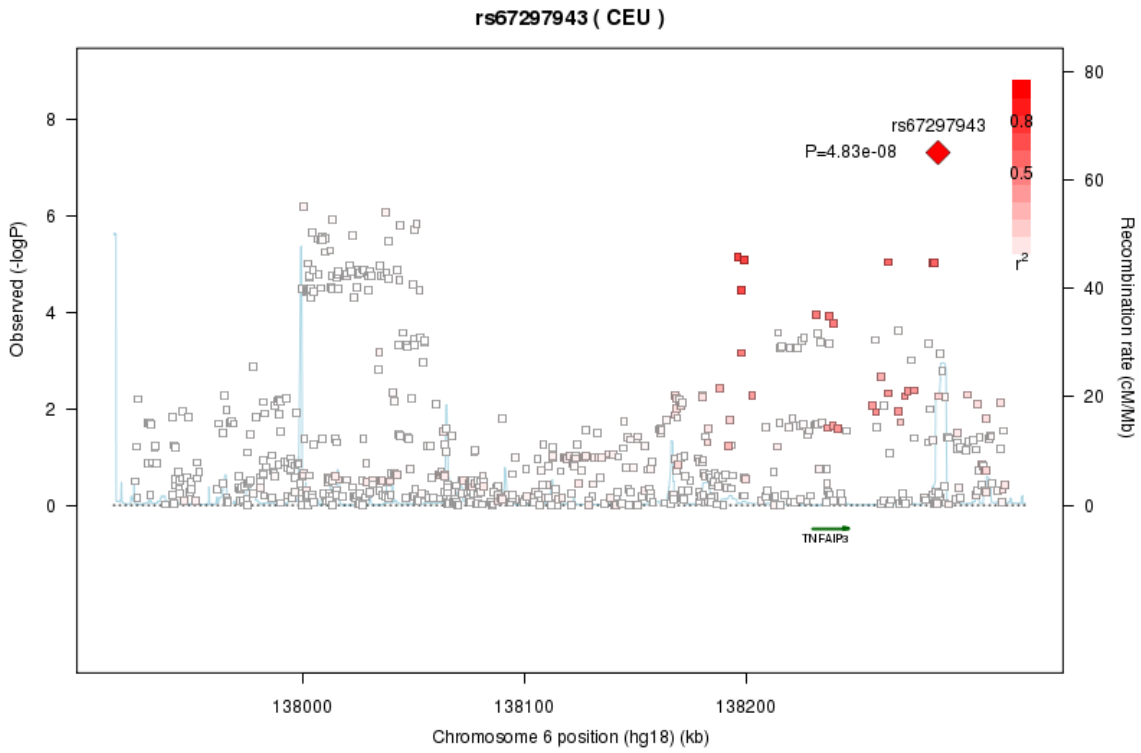
B



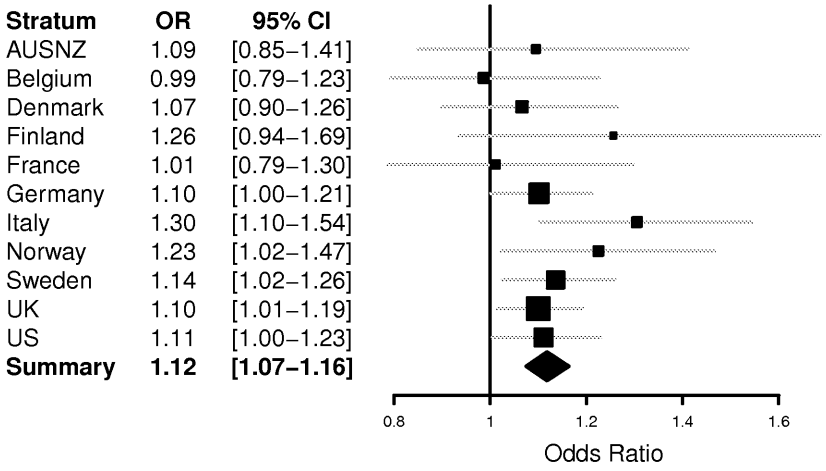
A) Regional Association and B) Forest Plot

Supplementary Figure 70. Discovery phase rs67297943.

A



B

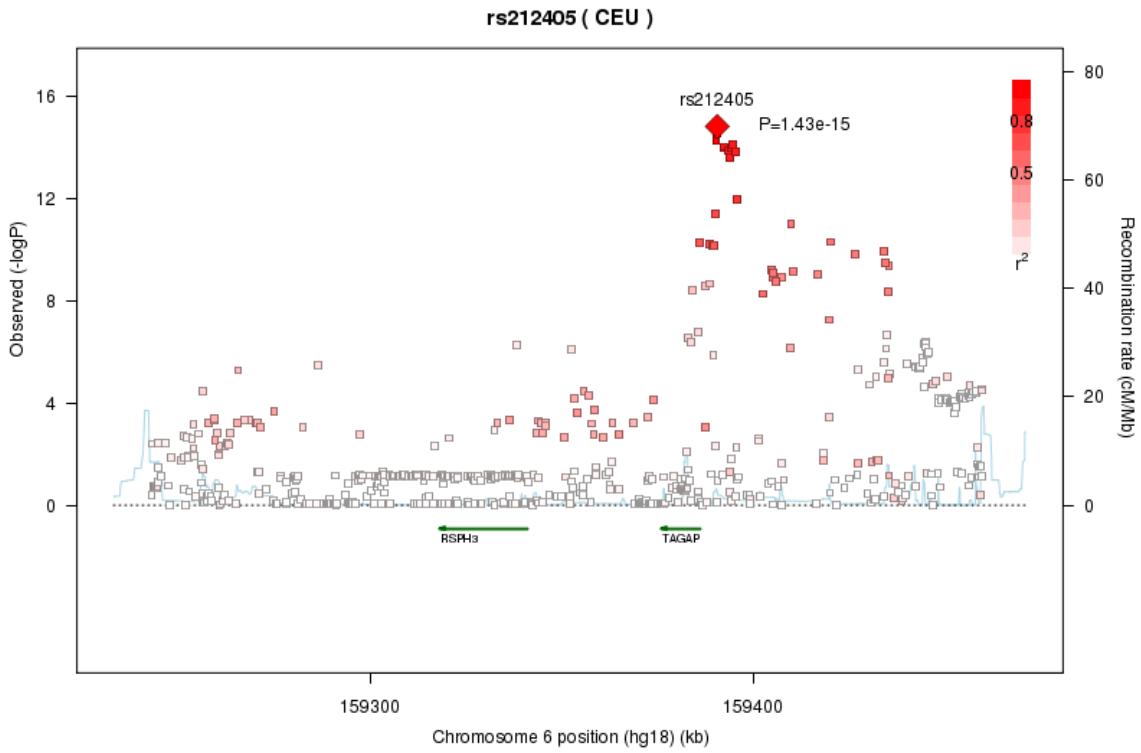


Stratum	OR	95% CI
AUSNZ	1.09	[0.85–1.41]
Belgium	0.99	[0.79–1.23]
Denmark	1.07	[0.90–1.26]
Finland	1.26	[0.94–1.69]
France	1.01	[0.79–1.30]
Germany	1.10	[1.00–1.21]
Italy	1.30	[1.10–1.54]
Norway	1.23	[1.02–1.47]
Sweden	1.14	[1.02–1.26]
UK	1.10	[1.01–1.19]
US	1.11	[1.00–1.23]
Summary	**1.12**	**[1.07–1.16]**

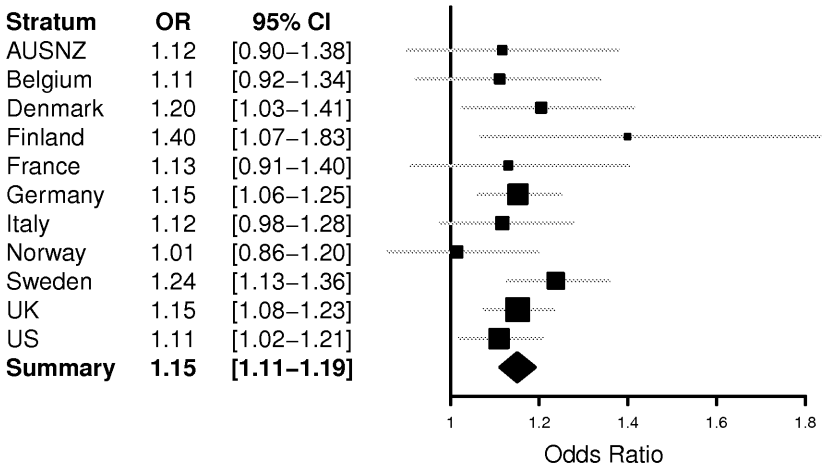
A) Regional Association and B) Forest Plot

Supplementary Figure 71. Discovery phase rs212405.

A



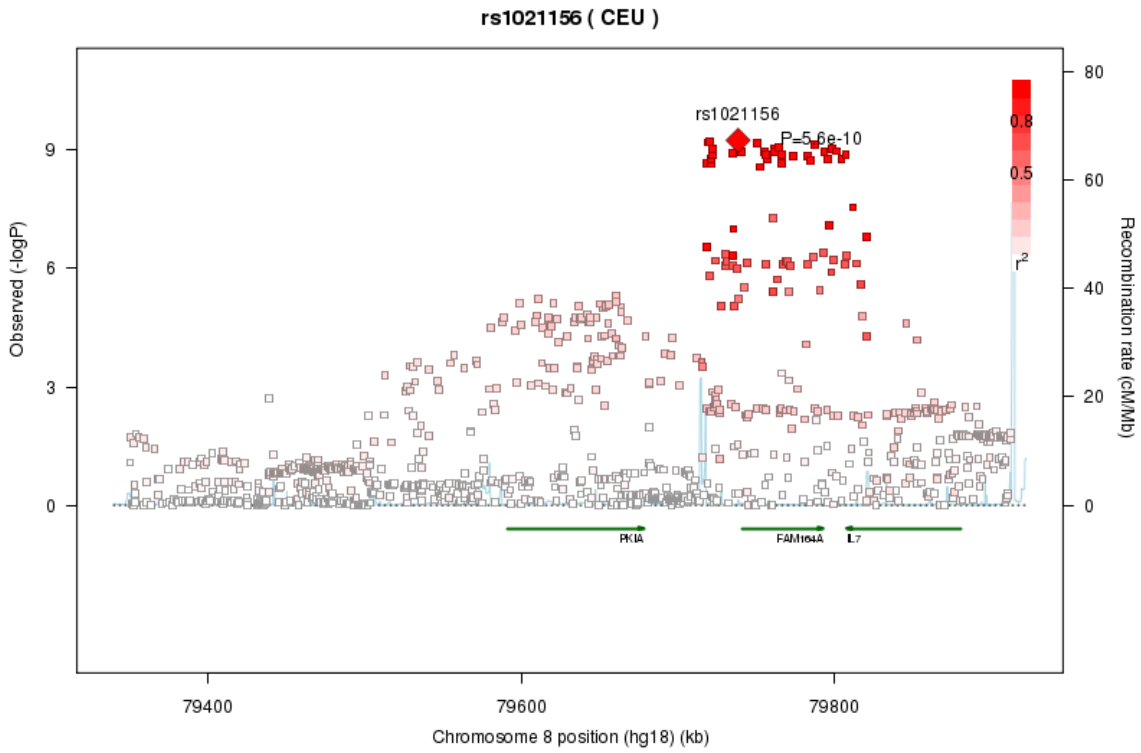
B



A) Regional Association and B) Forest Plot

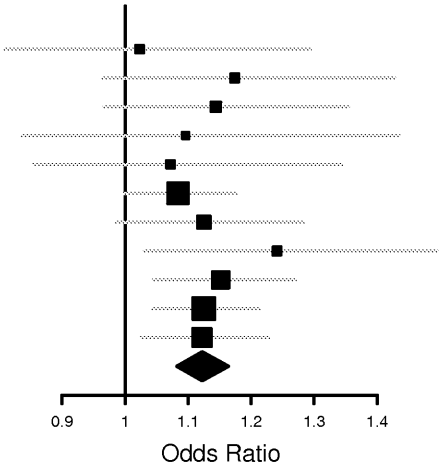
Supplementary Figure 72. Discovery phase rs1021156.

A



B

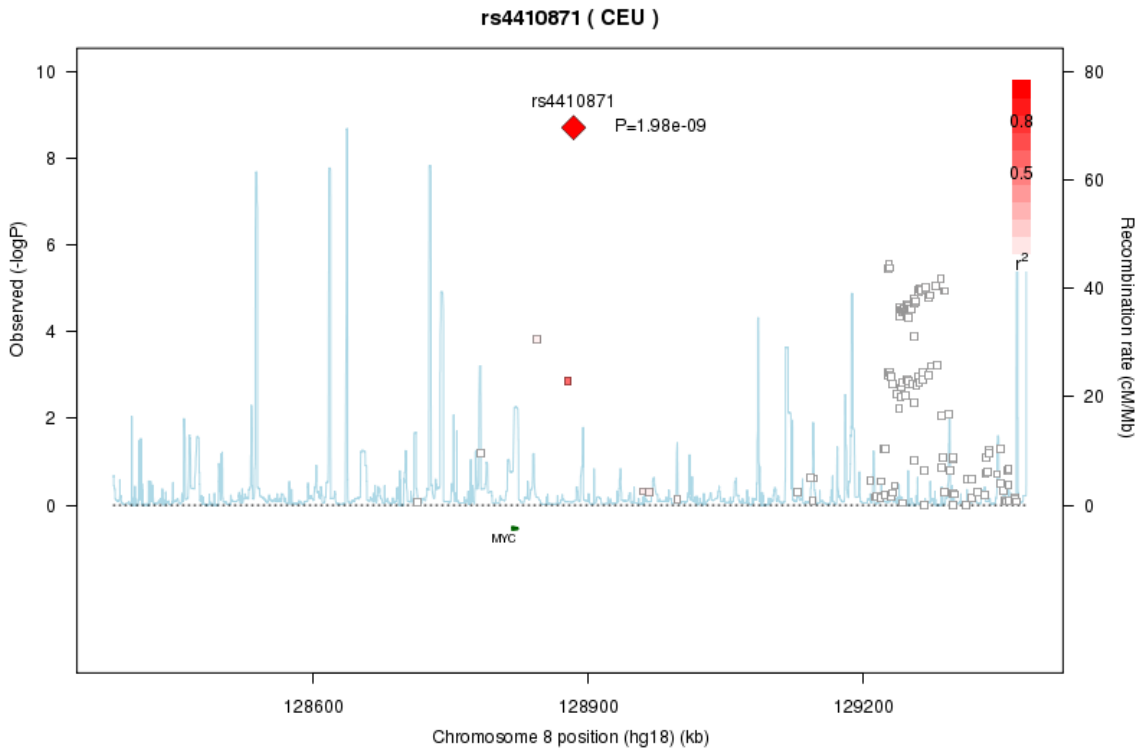
Stratum	OR	95% CI
AUSNZ	1.02	[0.81–1.29]
Belgium	1.17	[0.97–1.43]
Denmark	1.14	[0.97–1.35]
Finland	1.10	[0.84–1.43]
France	1.07	[0.85–1.34]
Germany	1.08	[1.00–1.18]
Italy	1.12	[0.99–1.28]
Norway	1.24	[1.03–1.49]
Sweden	1.15	[1.05–1.27]
UK	1.12	[1.04–1.21]
US	1.12	[1.03–1.23]
Summary	**1.12**	**[1.08–1.16]**



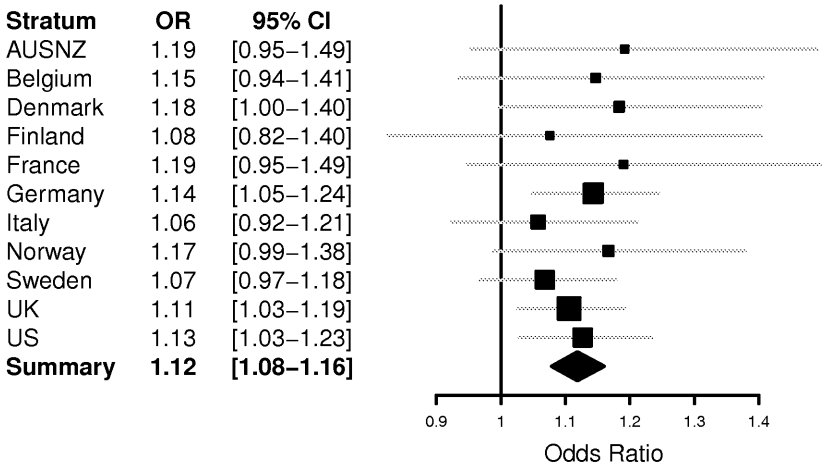
A) Regional Association and B) Forest Plot

Supplementary Figure 73. Discovery phase rs4410871.

A



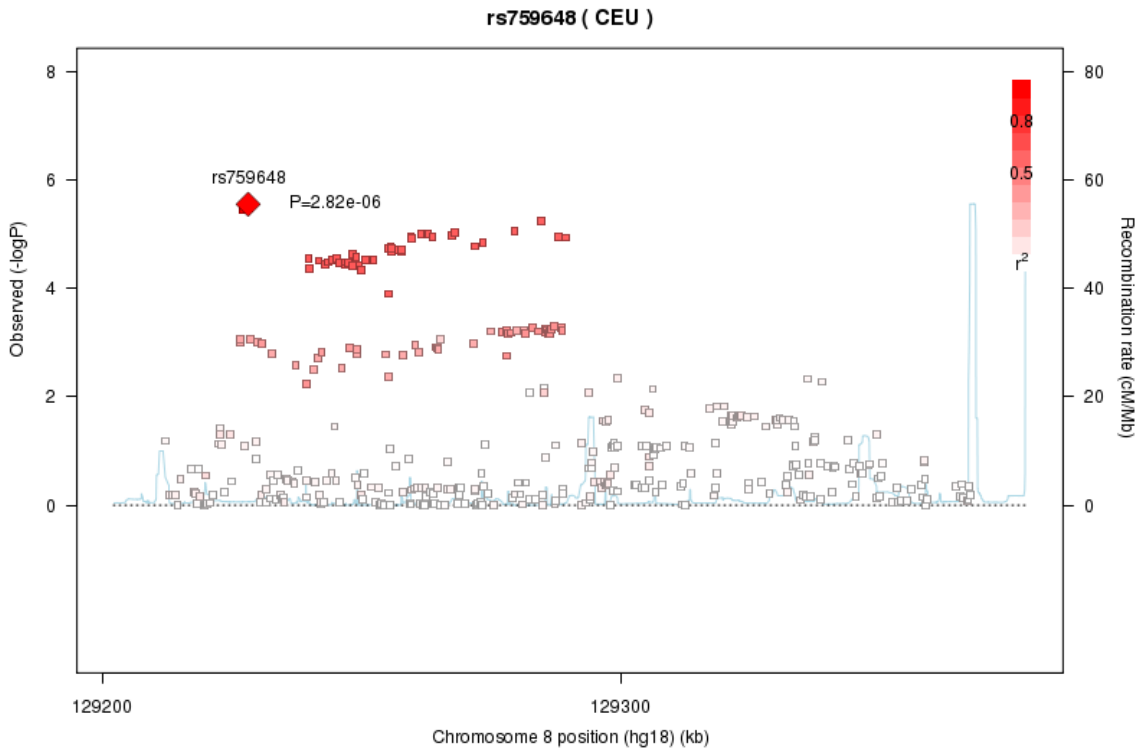
B



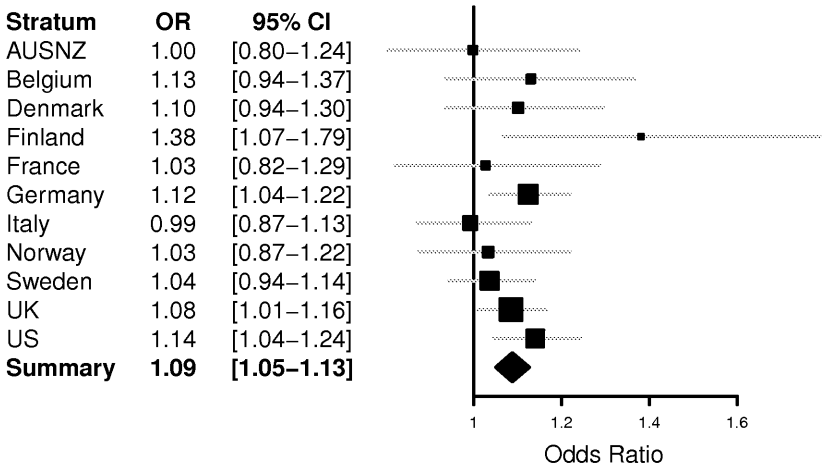
A) Regional Association and B) Forest Plot

Supplementary Figure 74. Discovery phase rs759648.

A



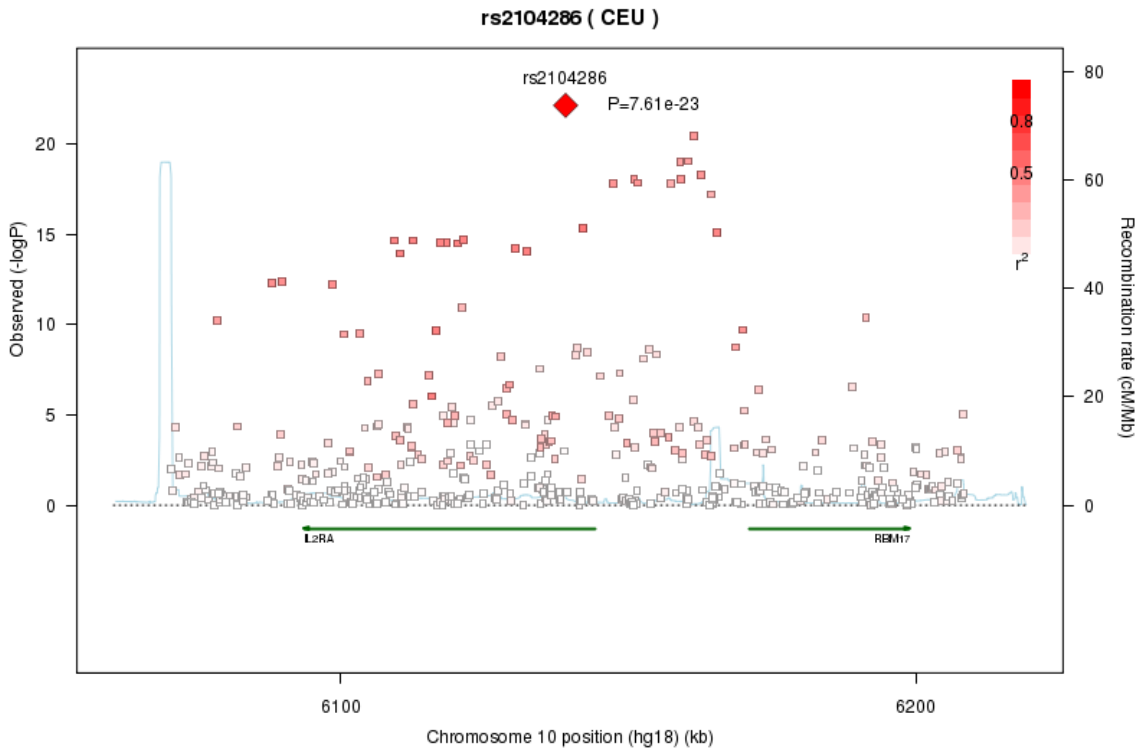
B



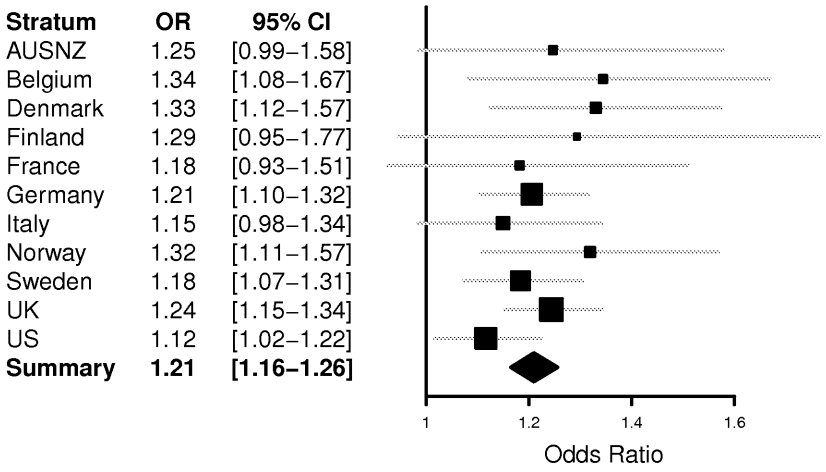
A) Regional Association and B) Forest Plot

Supplementary Figure 75: Discovery phase rs2104286.

A



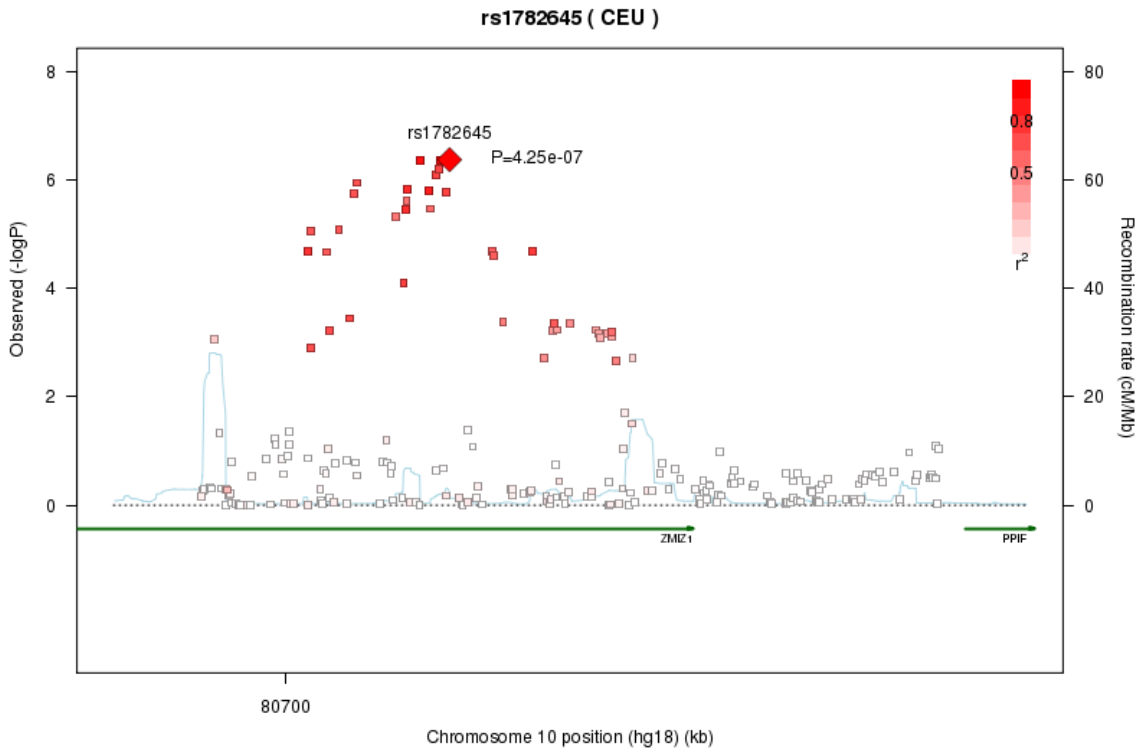
B



A) Regional Association and B) Forest Plot

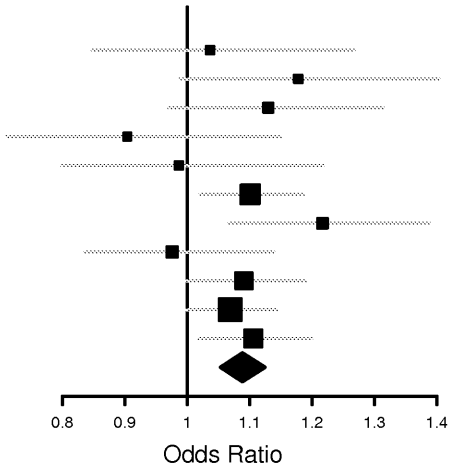
Supplementary Figure 76: Discovery phase rs1782645.

A



B

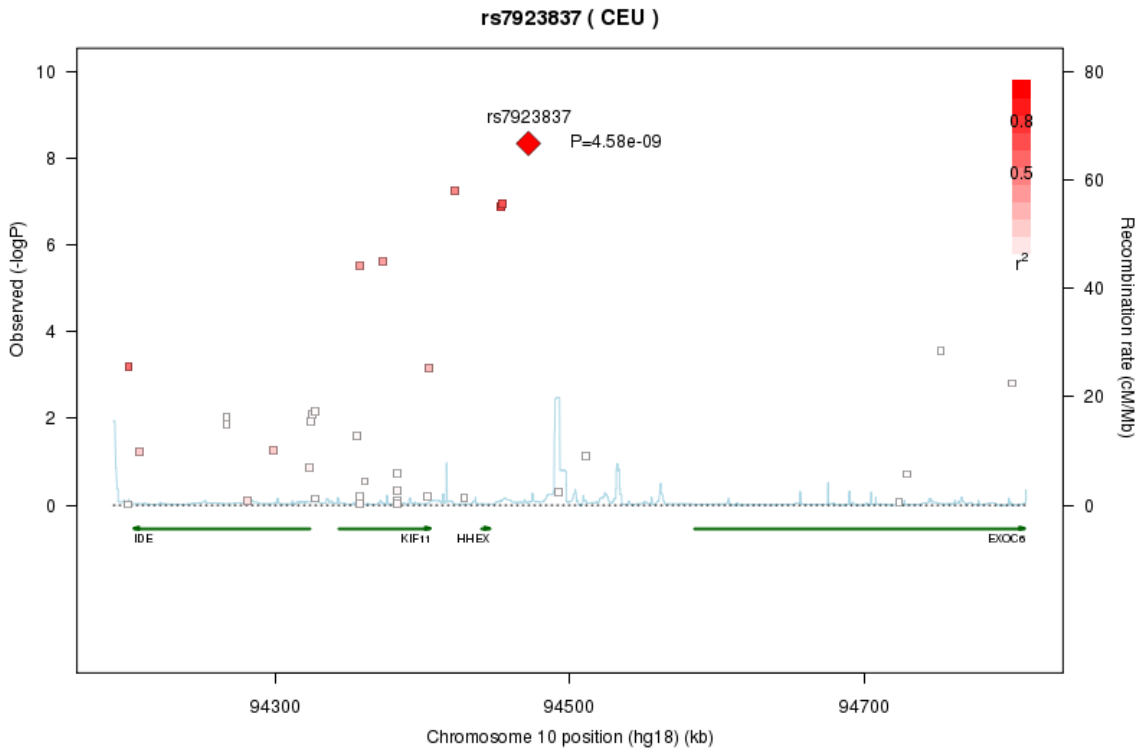
Stratum	OR	95% CI
AUSNZ	1.04	[0.85–1.27]
Belgium	1.18	[0.99–1.40]
Denmark	1.13	[0.97–1.31]
Finland	0.90	[0.71–1.15]
France	0.99	[0.80–1.22]
Germany	1.10	[1.02–1.19]
Italy	1.22	[1.07–1.39]
Norway	0.98	[0.84–1.14]
Sweden	1.09	[1.00–1.19]
UK	1.07	[1.00–1.14]
US	1.11	[1.02–1.20]
Summary	**1.09**	**[1.05–1.13]**



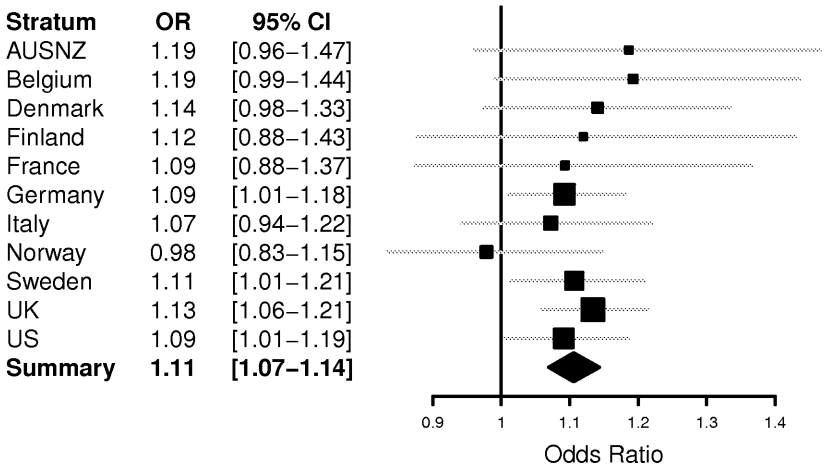
A) Regional Association and B) Forest Plot

Supplementary Figure 77. Discovery phase rs7923837.

A



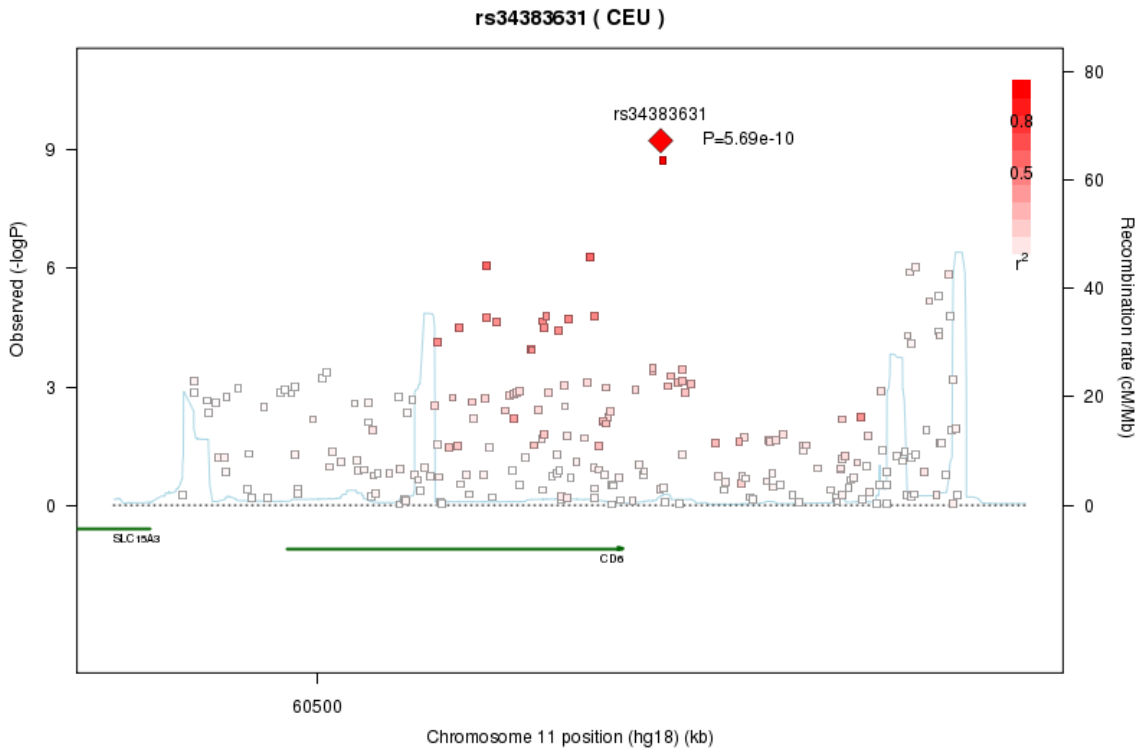
B



A) Regional Association and B) Forest Plot

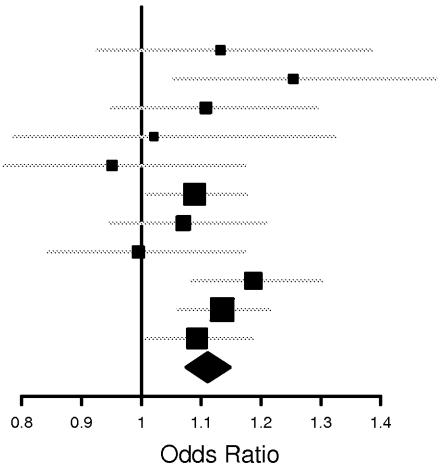
Supplementary Figure 78. Discovery phase rs34383631.

A



B

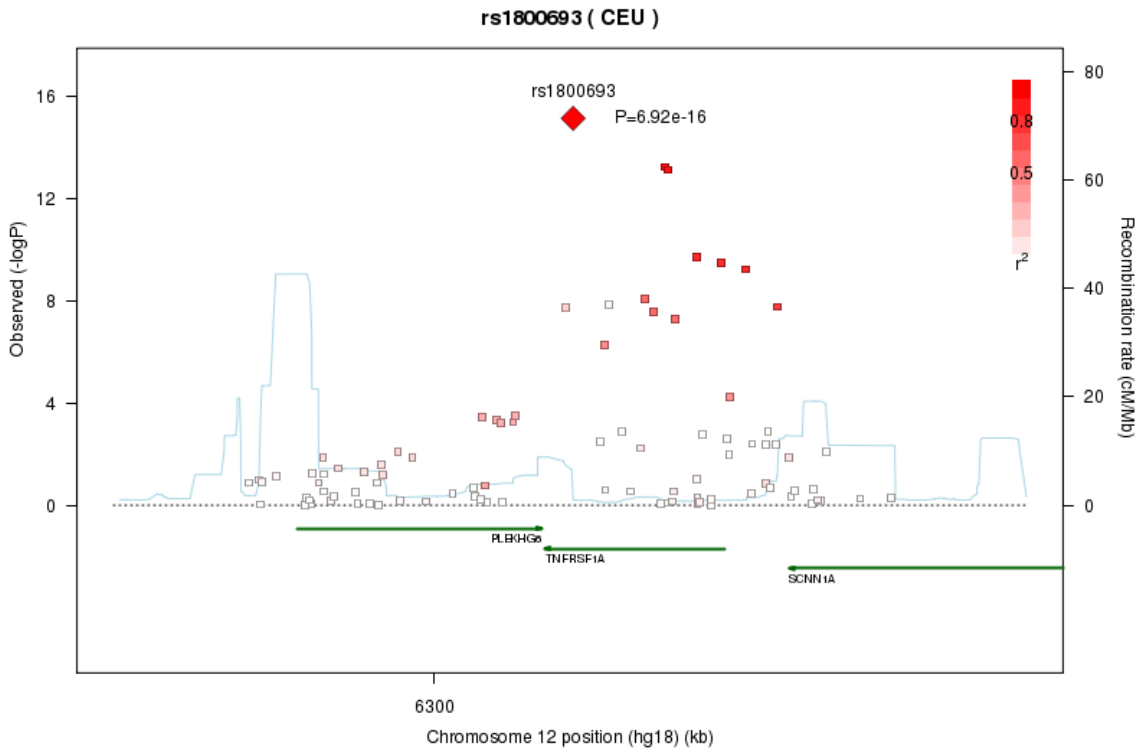
Stratum	OR	95% CI
AUSNZ	1.13	[0.93–1.38]
Belgium	1.25	[1.05–1.49]
Denmark	1.11	[0.95–1.29]
Finland	1.02	[0.79–1.32]
France	0.95	[0.77–1.17]
Germany	1.09	[1.01–1.18]
Italy	1.07	[0.95–1.21]
Norway	0.99	[0.84–1.17]
Sweden	1.19	[1.08–1.30]
UK	1.14	[1.06–1.21]
US	1.09	[1.01–1.19]
Summary	**1.11**	**[1.07–1.15]**



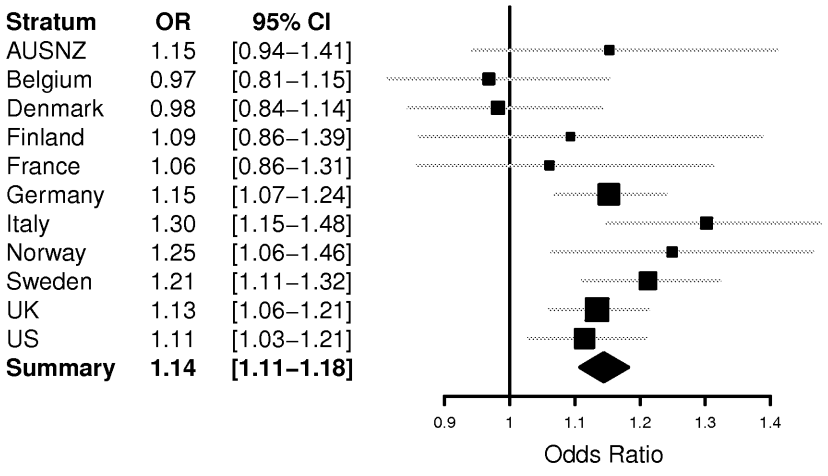
A) Regional Association and B) Forest Plot

Supplementary Figure 79. Discovery phase rs1800693.

A



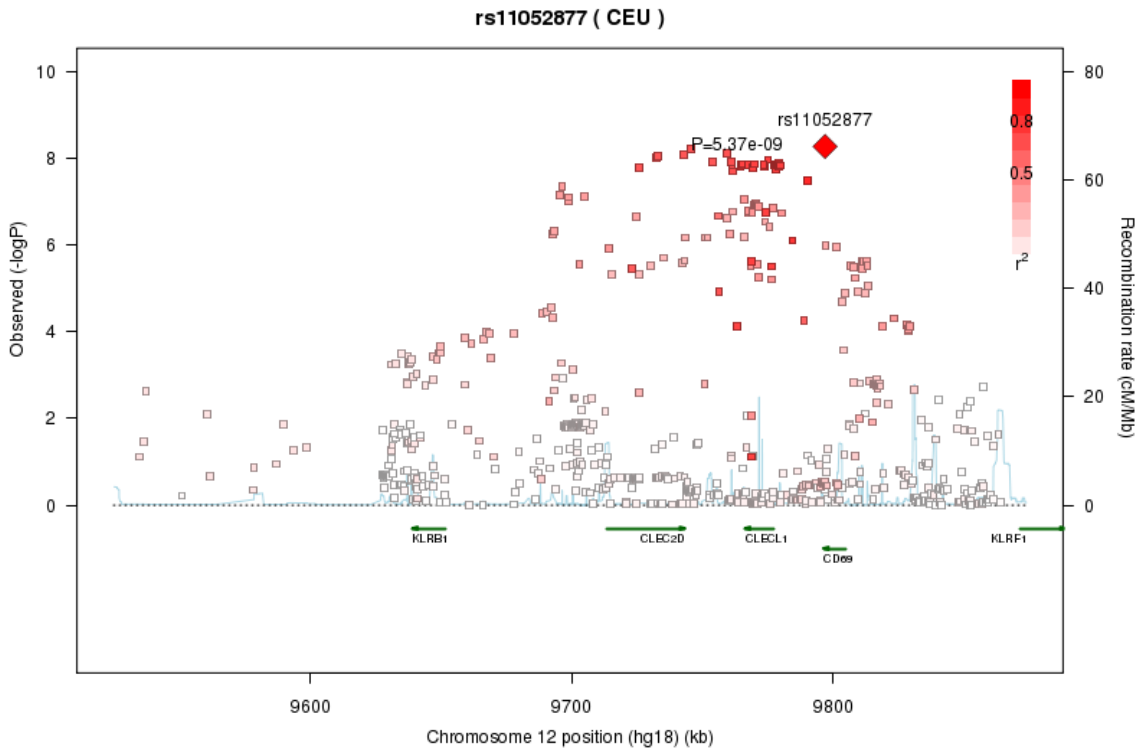
B



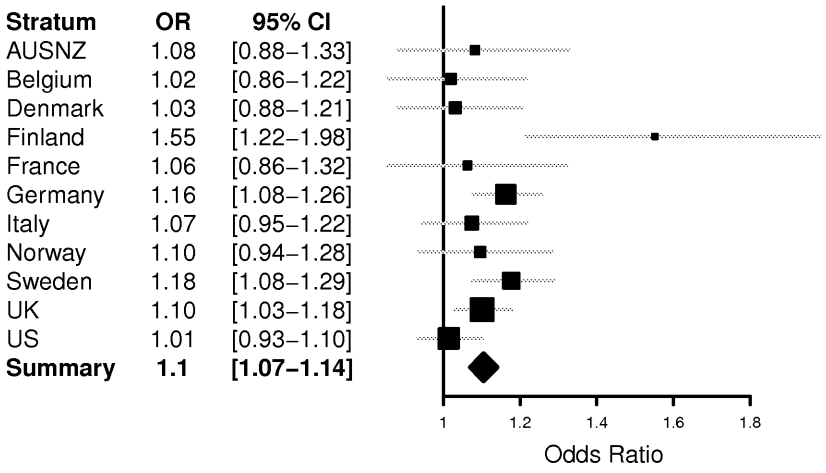
A) Regional Association and B) Forest Plot for rs1800693

Supplementary Figure 80. Discovery phase rs11052877.

A



B

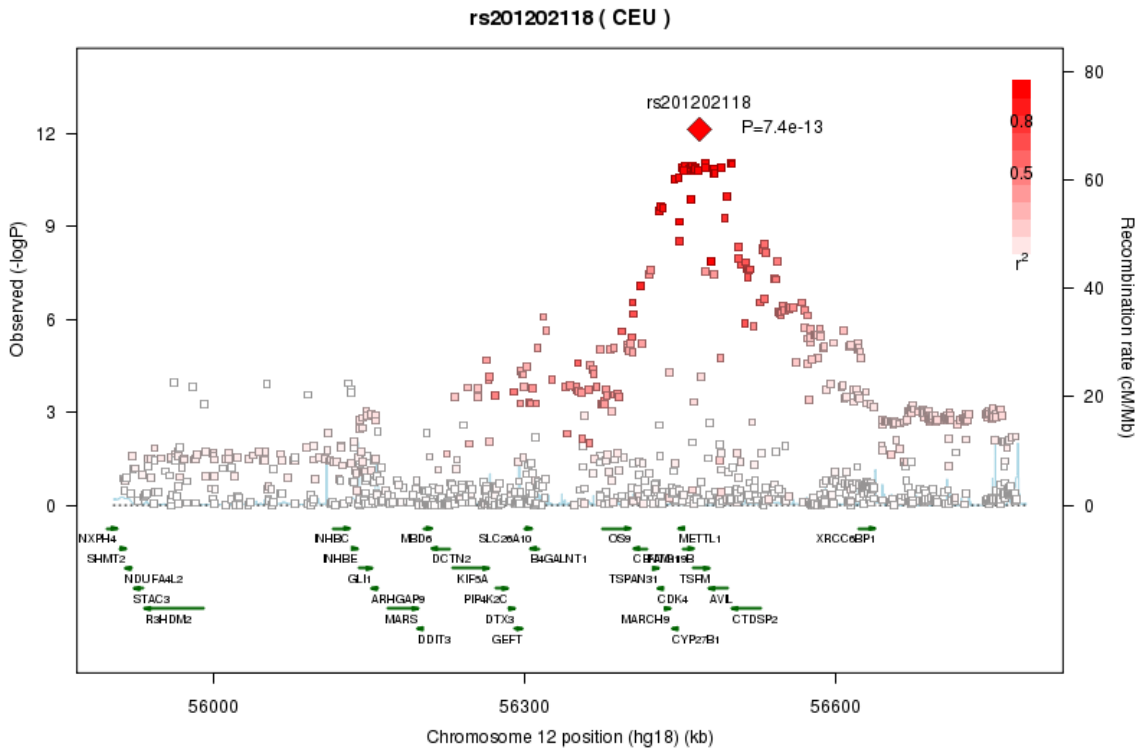


Stratum	OR	95% CI
AUSNZ	1.08	[0.88–1.33]
Belgium	1.02	[0.86–1.22]
Denmark	1.03	[0.88–1.21]
Finland	1.55	[1.22–1.98]
France	1.06	[0.86–1.32]
Germany	1.16	[1.08–1.26]
Italy	1.07	[0.95–1.22]
Norway	1.10	[0.94–1.28]
Sweden	1.18	[1.08–1.29]
UK	1.10	[1.03–1.18]
US	1.01	[0.93–1.10]
Summary	**1.1**	**[1.07–1.14]**

A) Regional Association and b) Forest Plot

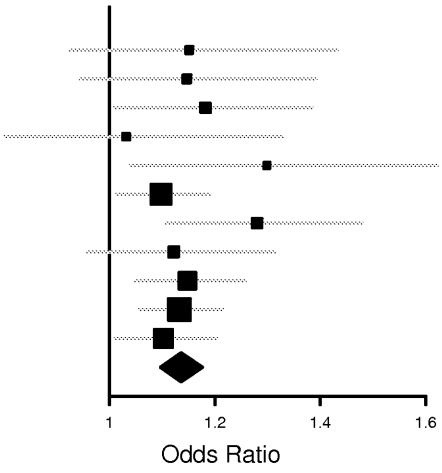
Supplementary Figure 81. Discovery phase rs201202118.

A



B

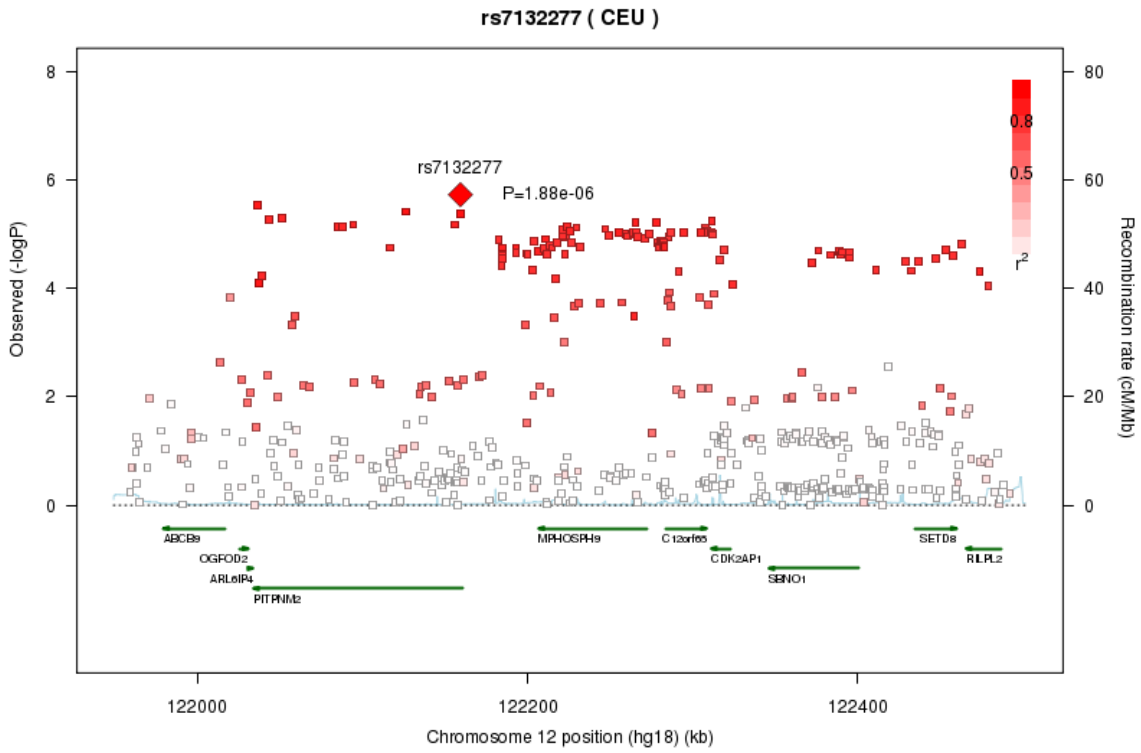
Stratum	OR	95% CI
AUSNZ	1.15	[0.93–1.43]
Belgium	1.15	[0.95–1.39]
Denmark	1.18	[1.01–1.38]
Finland	1.03	[0.80–1.33]
France	1.30	[1.04–1.62]
Germany	1.10	[1.01–1.19]
Italy	1.28	[1.11–1.48]
Norway	1.12	[0.96–1.31]
Sweden	1.15	[1.05–1.26]
UK	1.13	[1.06–1.22]
US	1.10	[1.01–1.20]
Summary	**1.14**	**[1.10–1.18]**



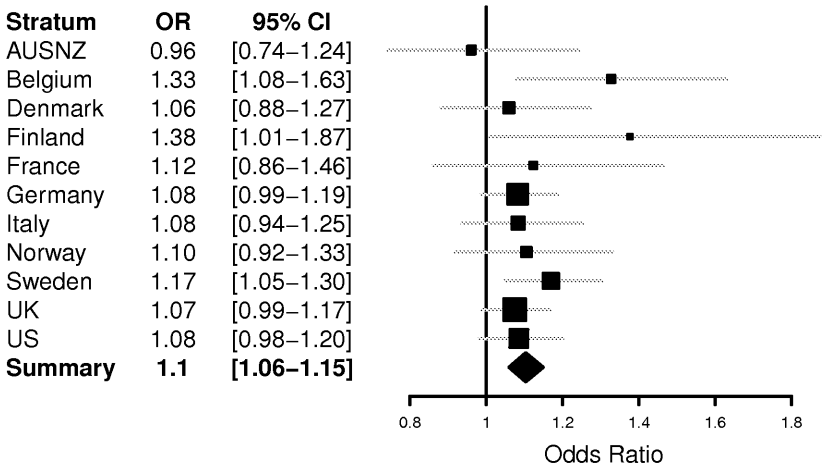
A) Regional Association and B) Forest Plot

Supplementary Figure 82. Discovery phase rs7132277.

A



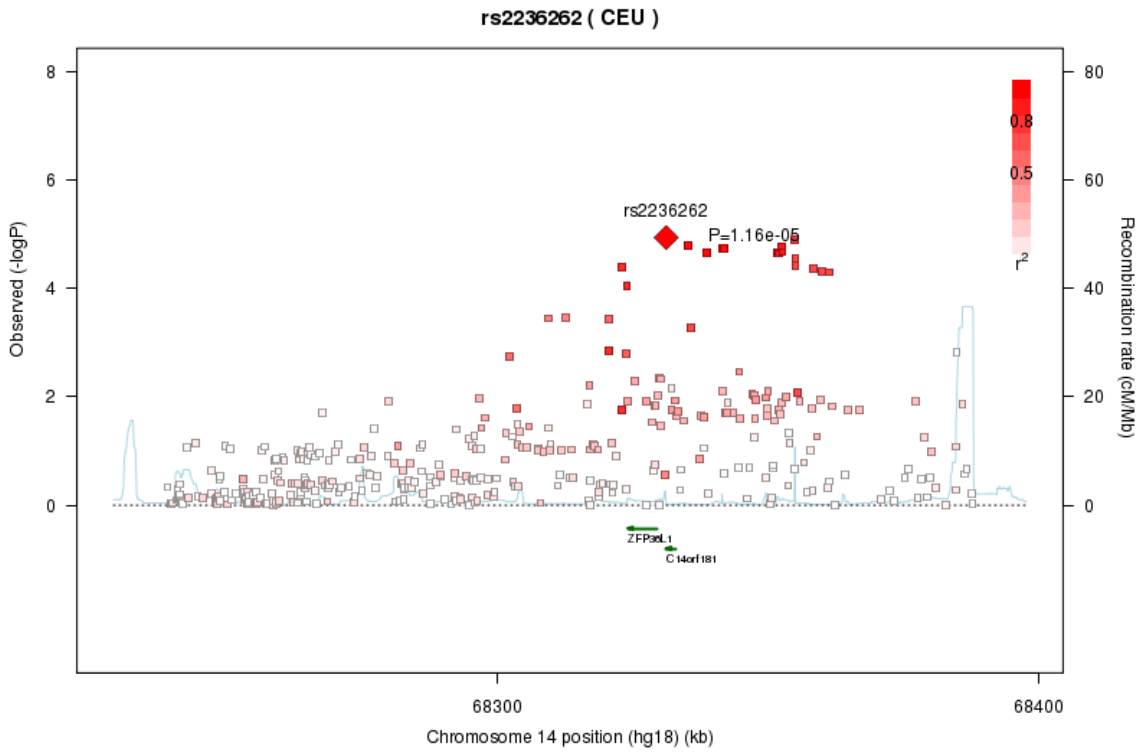
B



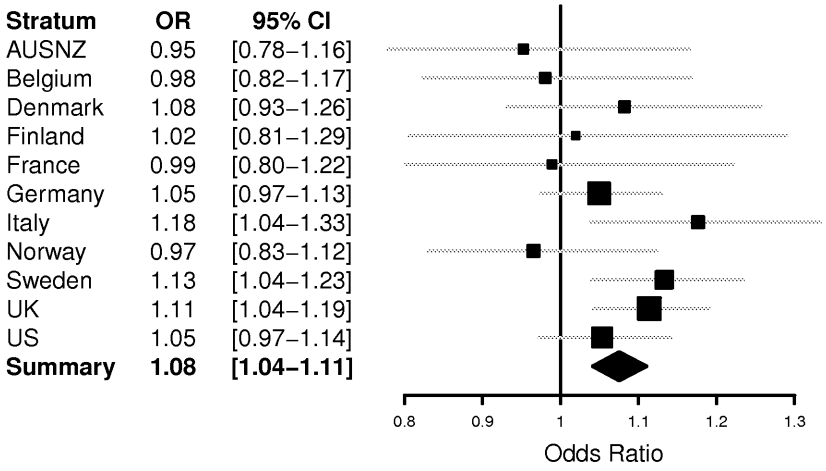
A) Regional Association and B) Forest Plot

Supplementary Figure 83. Discovery phase rs2236262.

A



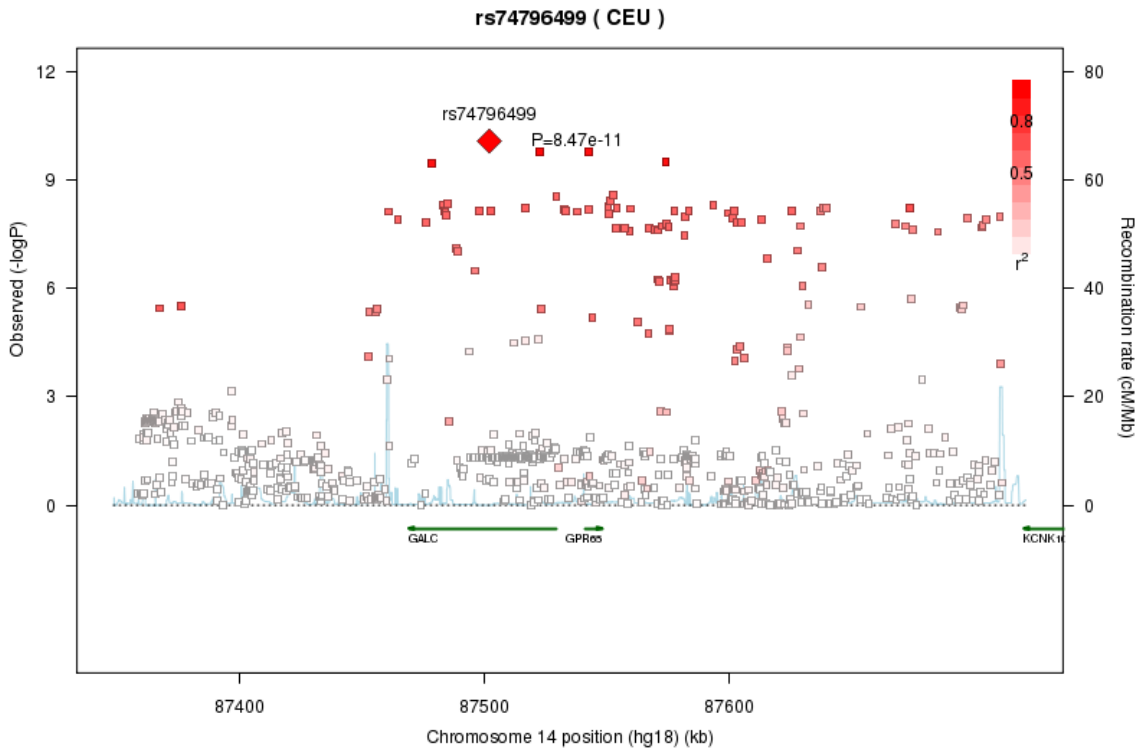
B



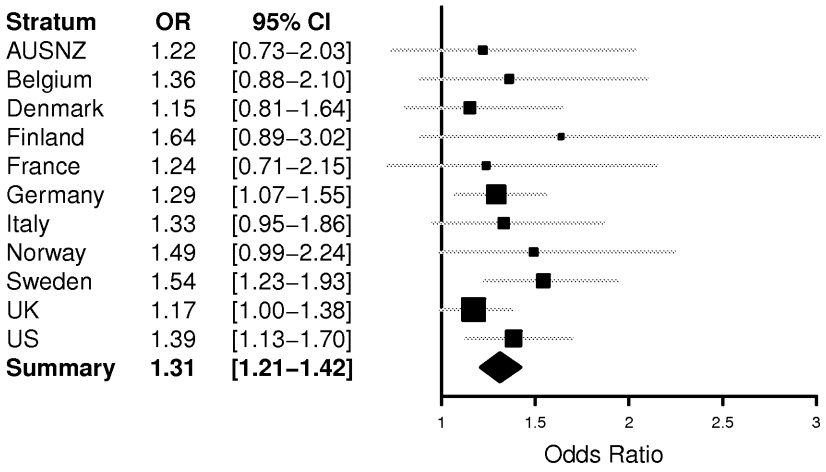
A) Regional Association and b) Forest Plot

Supplementary Figure 84. Discovery phase rs74796499.

A



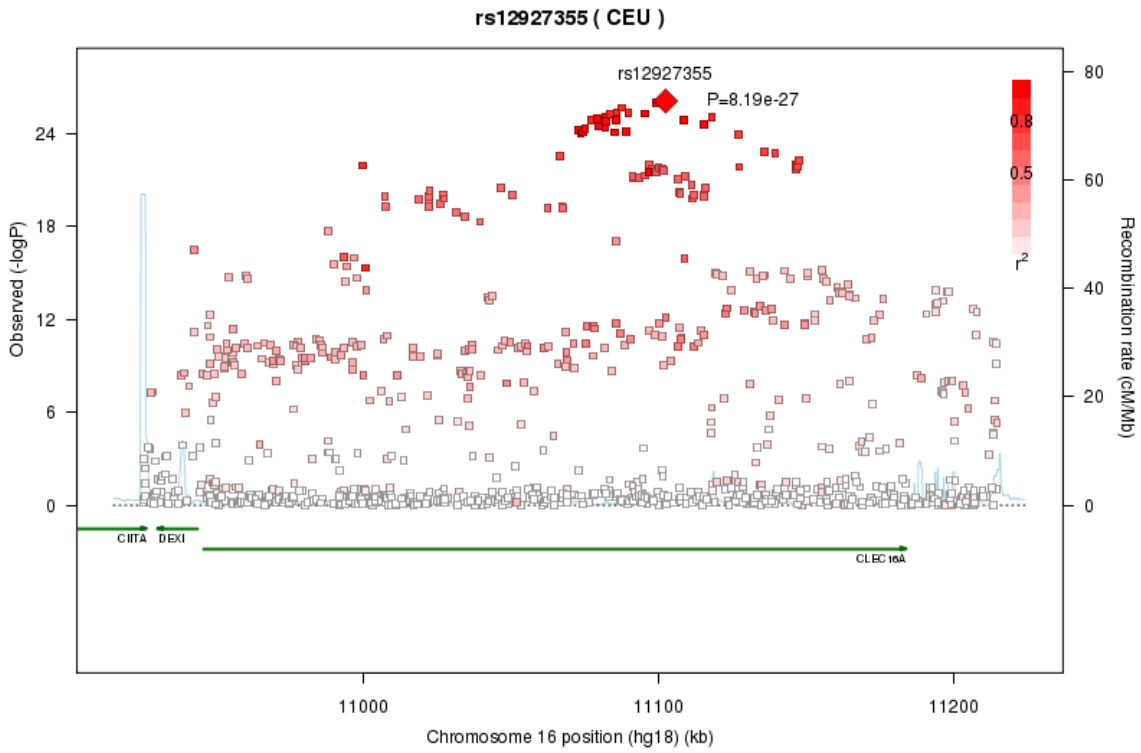
B



A) Regional Association and B) Forest Plot

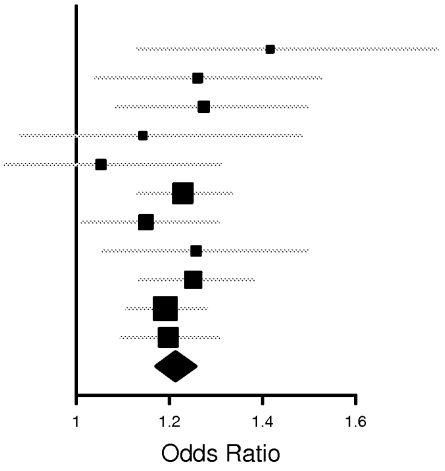
Supplementary Figure 85. Discovery phase rs12927355.

A



B

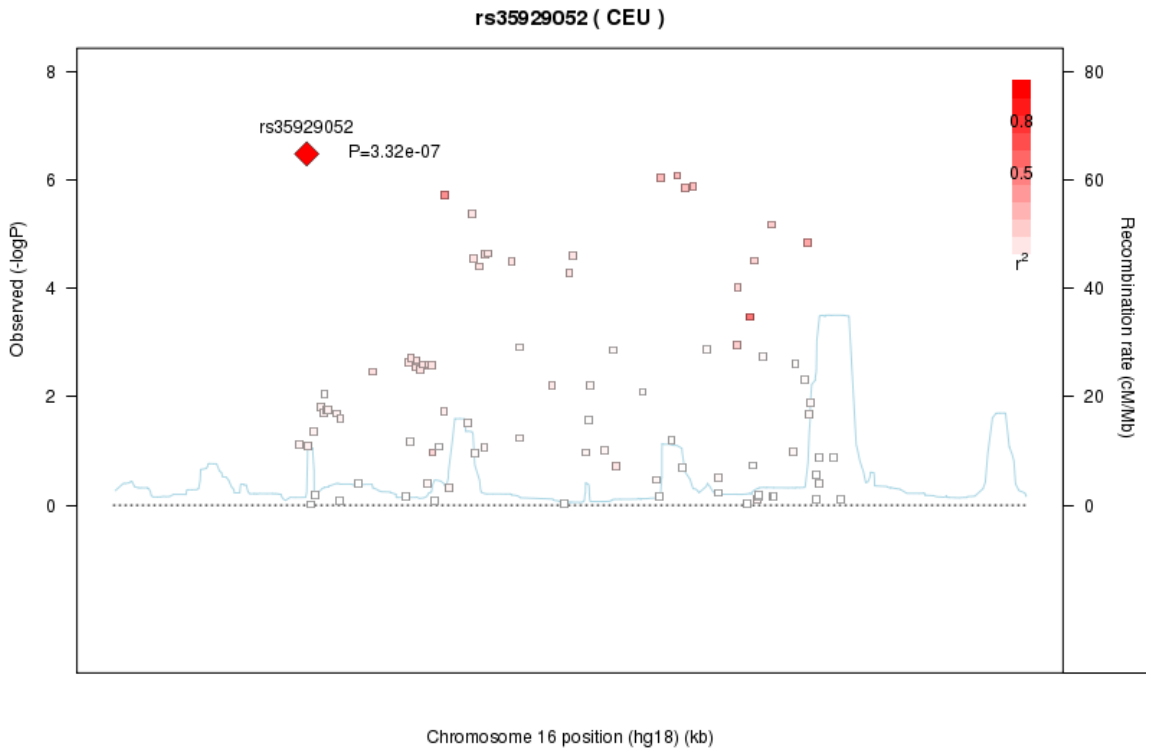
Stratum	OR	95% CI
AUSNZ	1.42	[1.13–1.77]
Belgium	1.26	[1.04–1.52]
Denmark	1.27	[1.09–1.50]
Finland	1.14	[0.88–1.48]
France	1.05	[0.85–1.31]
Germany	1.23	[1.13–1.33]
Italy	1.15	[1.01–1.30]
Norway	1.26	[1.06–1.50]
Sweden	1.25	[1.14–1.38]
UK	1.19	[1.11–1.28]
US	1.20	[1.10–1.31]
Summary	**1.21**	**[1.17–1.26]**



A) Regional Association and B) Forest Plot

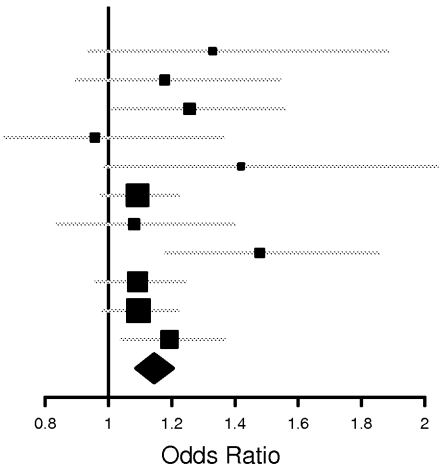
Supplementary Figure 86. Discovery phase rs35929052.

A



B

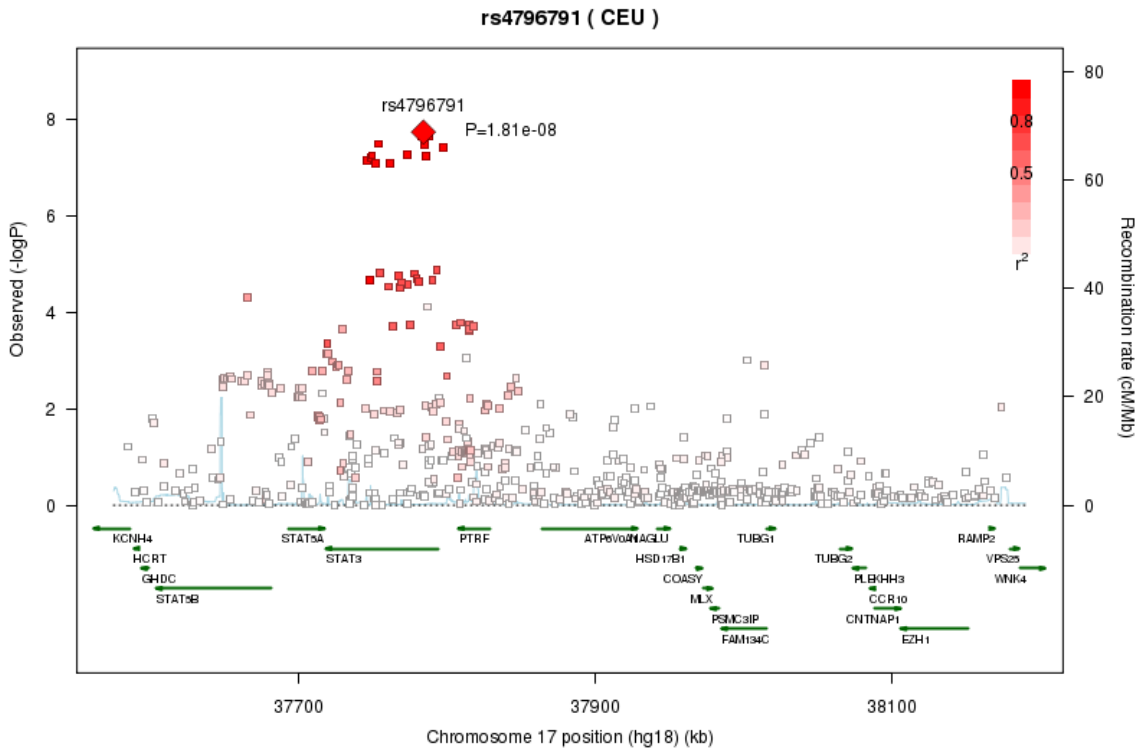
Stratum	OR	95% CI
AUSNZ	1.33	[0.94–1.88]
Belgium	1.18	[0.90–1.54]
Denmark	1.26	[1.01–1.56]
Finland	0.96	[0.67–1.36]
France	1.42	[0.99–2.04]
Germany	1.09	[0.97–1.22]
Italy	1.08	[0.84–1.40]
Norway	1.48	[1.18–1.85]
Sweden	1.09	[0.96–1.24]
UK	1.09	[0.98–1.22]
US	1.19	[1.04–1.37]
Summary	**1.14**	**[1.09–1.20]**



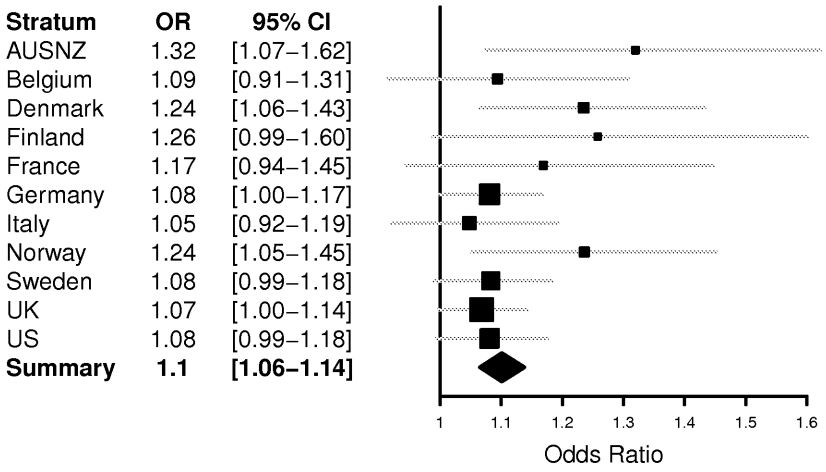
A) Regional Association and B) Forest Plot

Supplementary Figure 87. Discovery phase rs4796791.

A



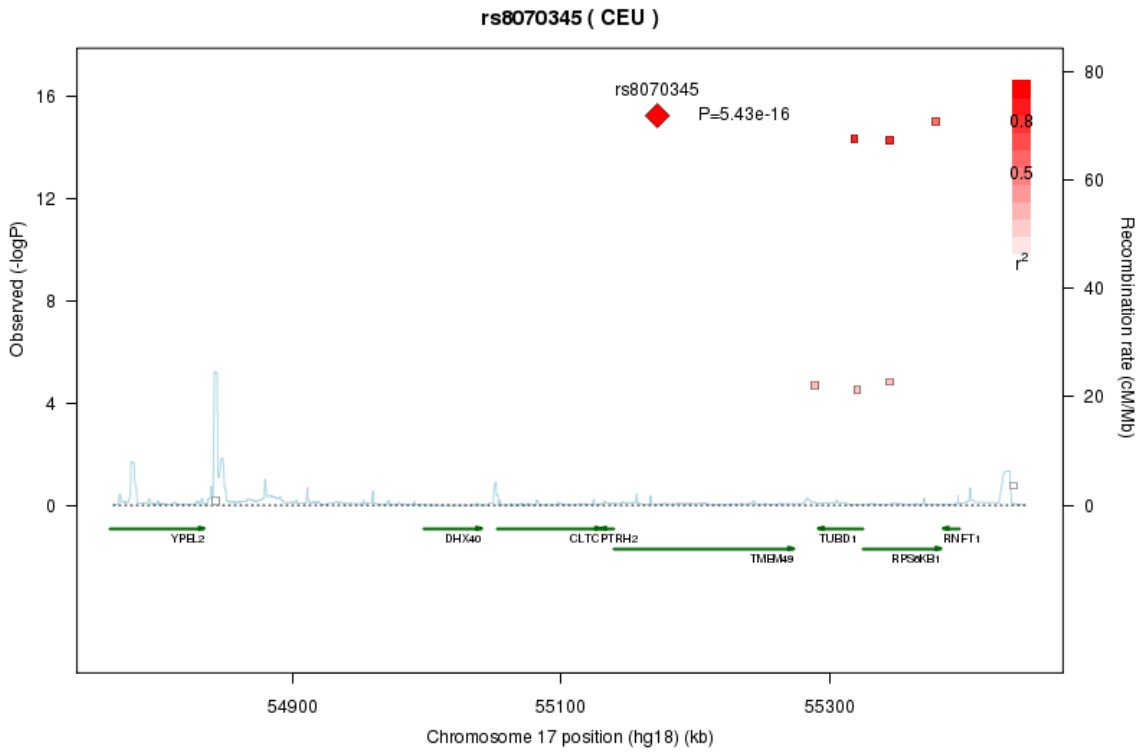
B



A) Regional Association and B) Forest Plot

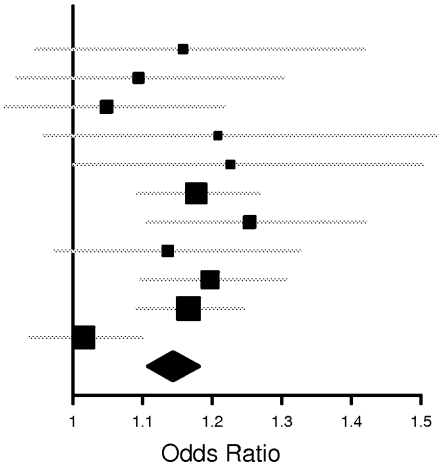
Supplementary Figure 88. Discovery phase rs8070345.

A



B

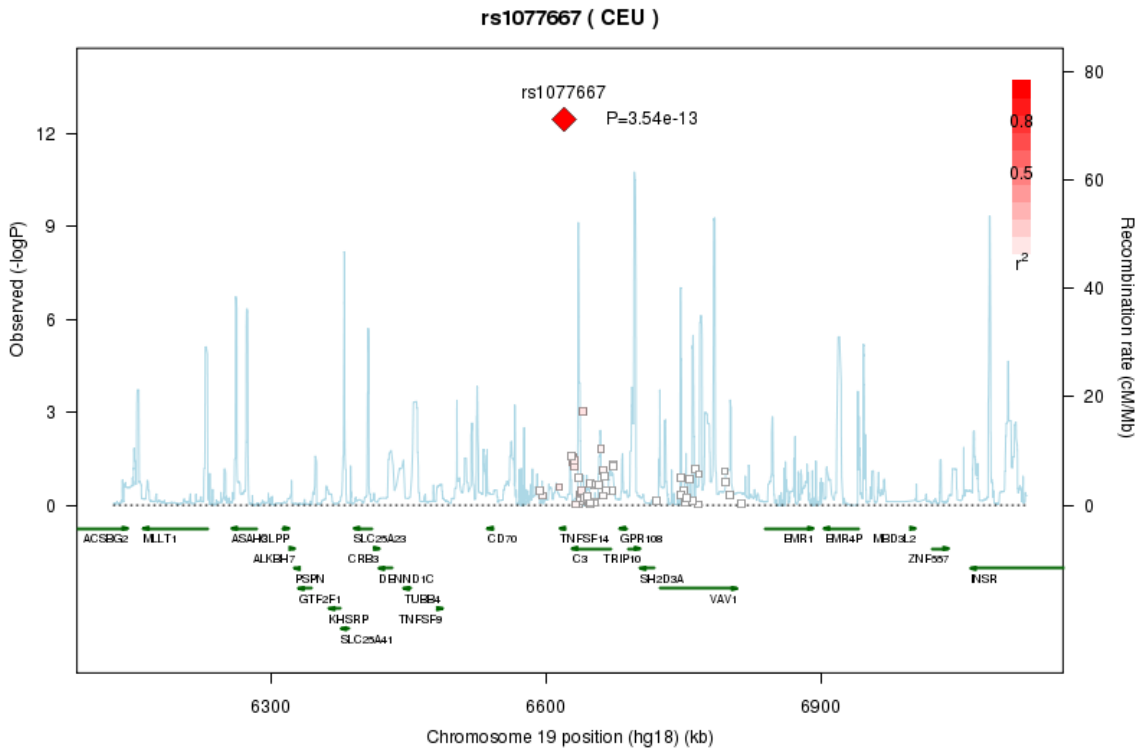
Stratum	OR	95% CI
AUSNZ	1.16	[0.95–1.42]
Belgium	1.09	[0.92–1.30]
Denmark	1.05	[0.90–1.22]
Finland	1.21	[0.96–1.52]
France	1.23	[1.00–1.50]
Germany	1.18	[1.09–1.27]
Italy	1.25	[1.11–1.42]
Norway	1.14	[0.97–1.33]
Sweden	1.20	[1.10–1.31]
UK	1.17	[1.09–1.25]
US	1.01	[0.94–1.10]
Summary	**1.14**	**[1.11–1.18]**



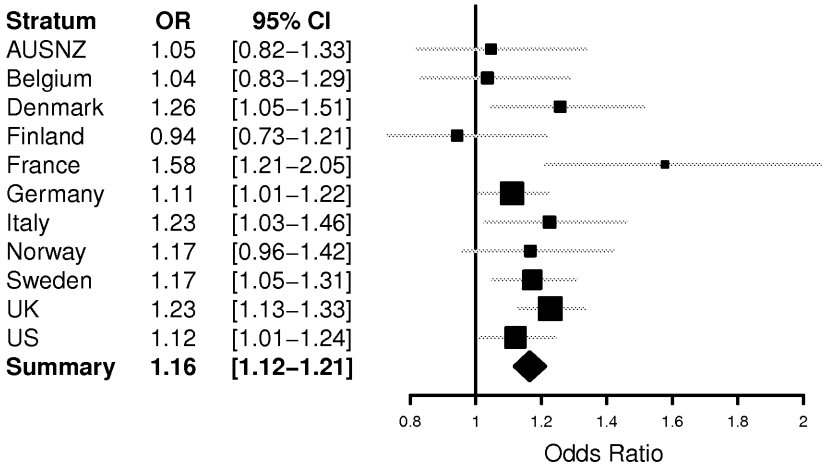
A) Regional Association and B) Forest Plot

Supplementary Figure 89. Discovery phase rs1077667.

A



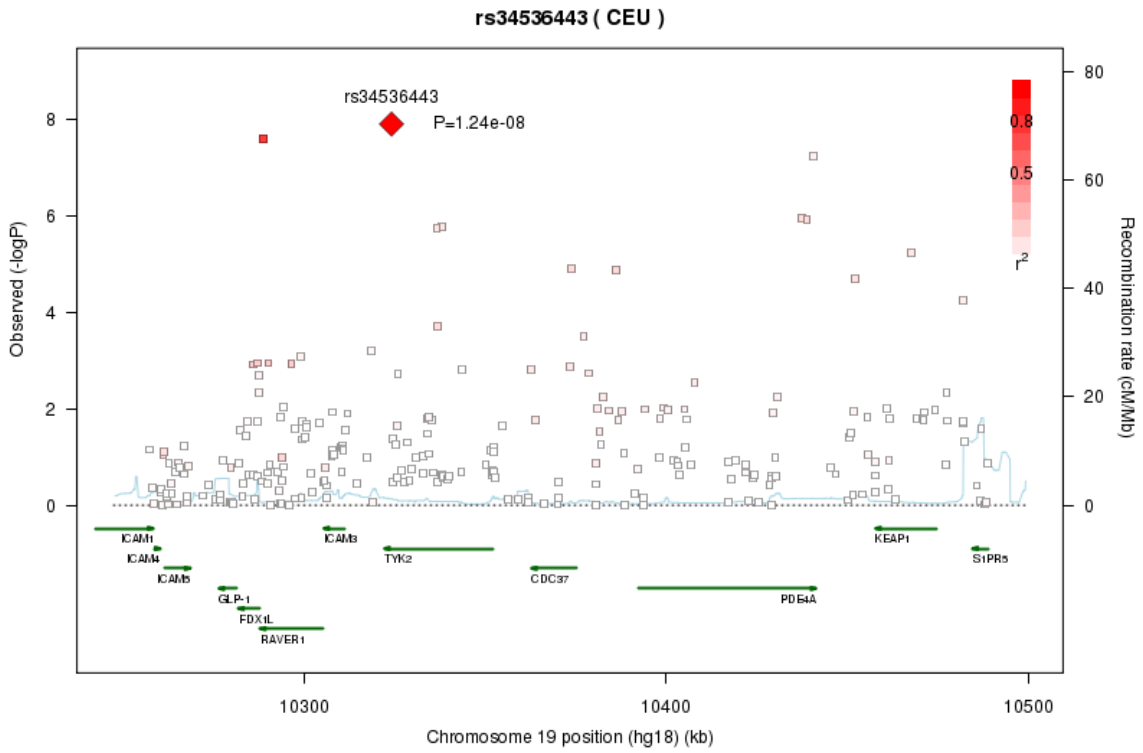
B



A) Regional Association and B) Forest Plot

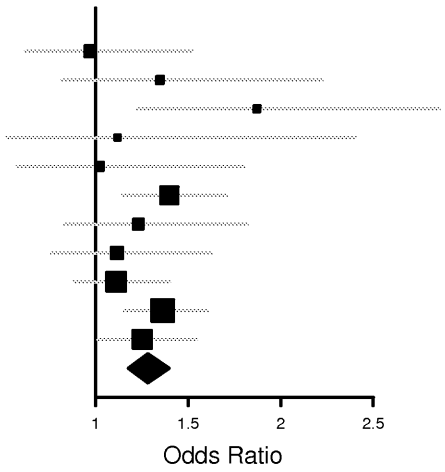
Supplementary Figure 90. Discovery phase rs34536443.

A



B

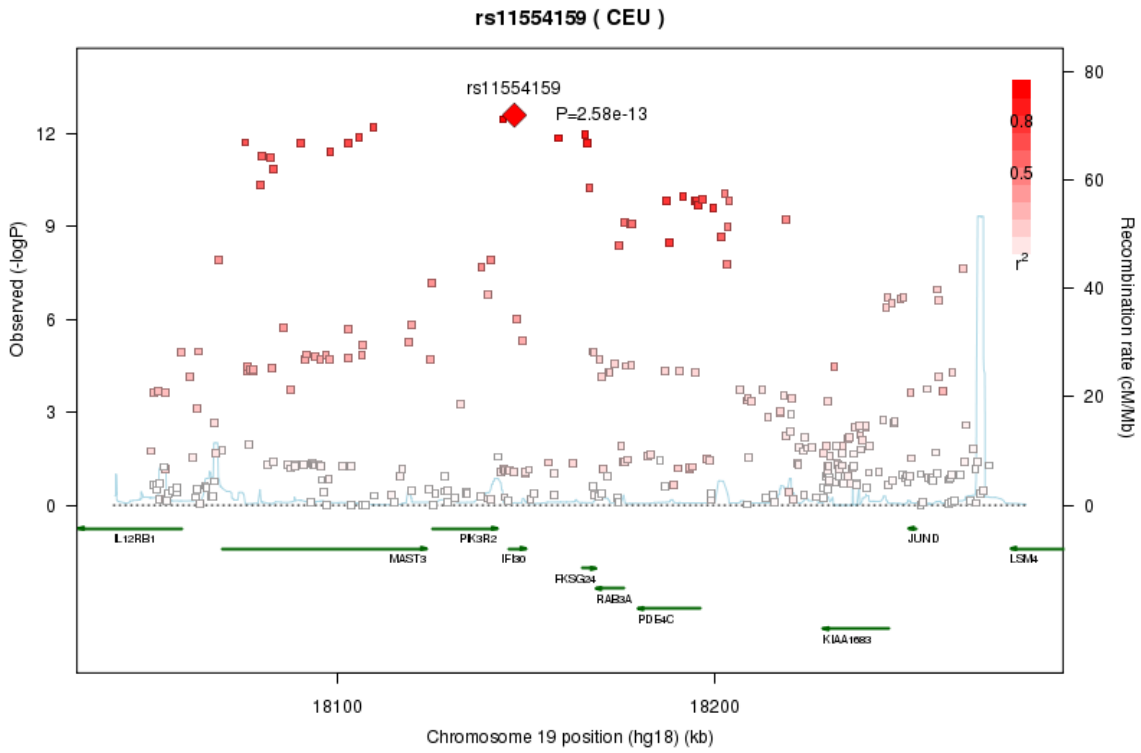
Stratum	OR	95% CI
AUSNZ	0.97	[0.62–1.52]
Belgium	1.35	[0.82–2.22]
Denmark	1.87	[1.23–2.85]
Finland	1.12	[0.52–2.40]
France	1.02	[0.58–1.80]
Germany	1.40	[1.14–1.71]
Italy	1.23	[0.83–1.82]
Norway	1.11	[0.76–1.63]
Sweden	1.11	[0.88–1.40]
UK	1.36	[1.16–1.60]
US	1.25	[1.01–1.54]
Summary	**1.28**	**[1.18–1.40]**



A) Regional Association and B) Forest Plot

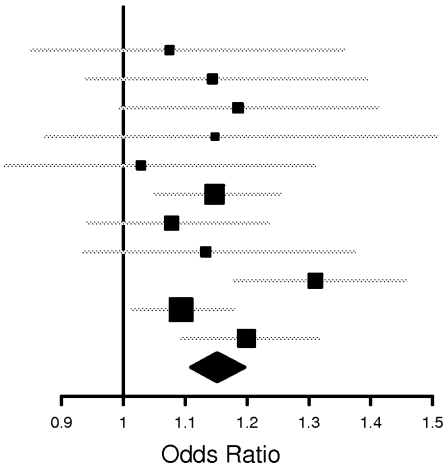
Supplementary Figure 91. Discovery phase rs11554159.

A



B

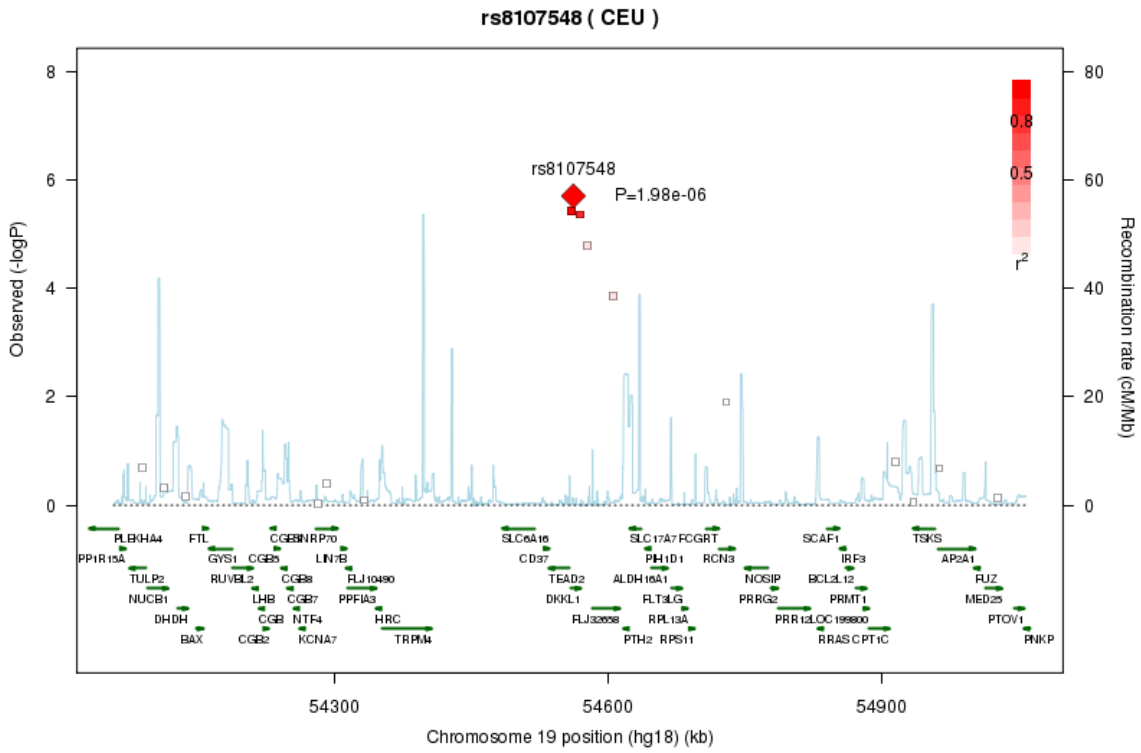
Stratum	OR	95% CI
AUSNZ	1.07	[0.85–1.36]
Belgium	1.14	[0.94–1.39]
Denmark	1.19	[1.00–1.41]
Finland	1.15	[0.87–1.51]
France	1.03	[0.81–1.31]
Germany	1.15	[1.05–1.25]
Italy	1.08	[0.94–1.23]
Norway	1.13	[0.94–1.37]
Sweden	1.31	[1.18–1.46]
UK	1.09	[1.01–1.18]
US	1.20	[1.09–1.32]
Summary	**1.15**	**[1.11–1.20]**



A) Regional Association and B) Forest Plot

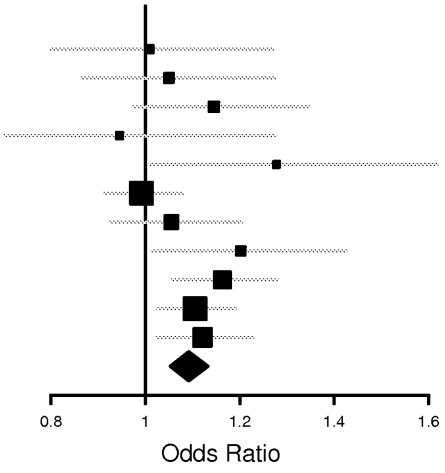
Supplementary Figure 92. Discovery phase rs8107548.

A



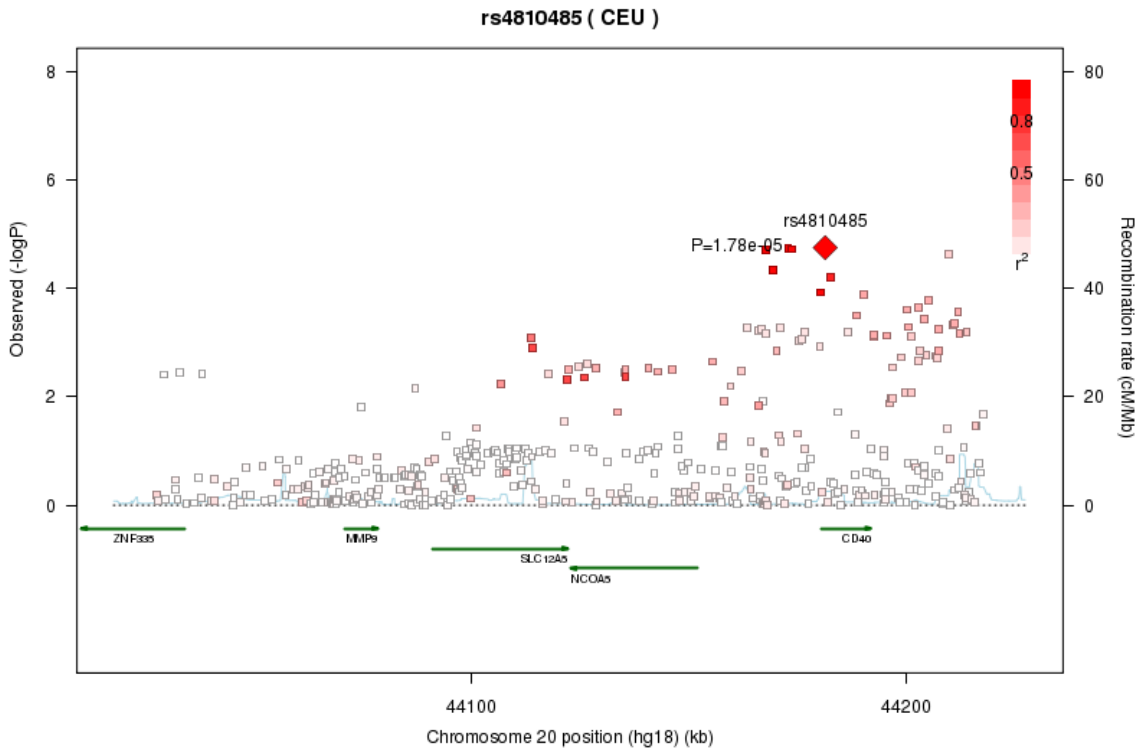
B

Stratum	OR	95% CI
AUSNZ	1.01	[0.80–1.27]
Belgium	1.05	[0.87–1.27]
Denmark	1.15	[0.97–1.35]
Finland	0.94	[0.70–1.27]
France	1.28	[1.01–1.62]
Germany	0.99	[0.91–1.08]
Italy	1.05	[0.93–1.20]
Norway	1.20	[1.01–1.43]
Sweden	1.16	[1.06–1.28]
UK	1.10	[1.03–1.19]
US	1.12	[1.02–1.23]
Summary	**1.09**	**[1.05–1.13]**



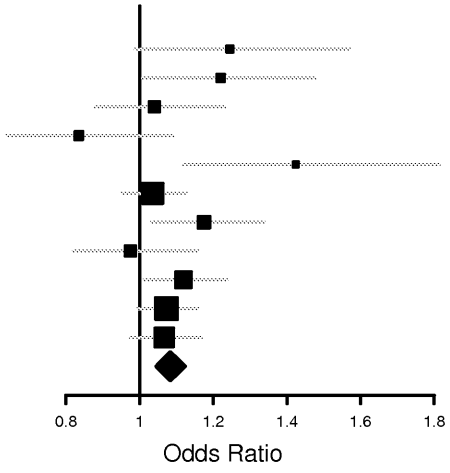
A) Regional Association and B) Forest Plot

Supplementary Figure 93. Discovery phase rs4810485.
A



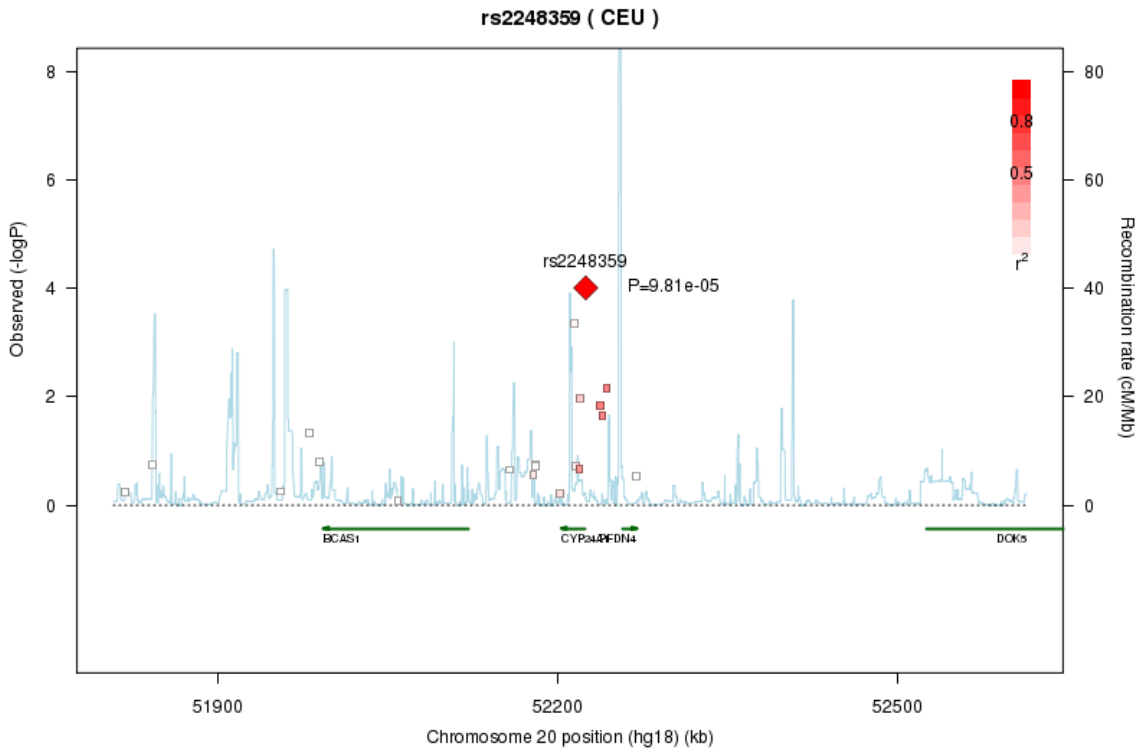
B

Stratum	OR	95% CI
AUSNZ	1.25	[0.99–1.57]
Belgium	1.22	[1.01–1.48]
Denmark	1.04	[0.88–1.23]
Finland	0.83	[0.64–1.09]
France	1.42	[1.12–1.81]
Germany	1.04	[0.95–1.13]
Italy	1.18	[1.03–1.34]
Norway	0.97	[0.82–1.16]
Sweden	1.12	[1.01–1.24]
UK	1.07	[0.99–1.16]
US	1.07	[0.98–1.17]
Summary	**1.08**	**[1.04–1.12]**

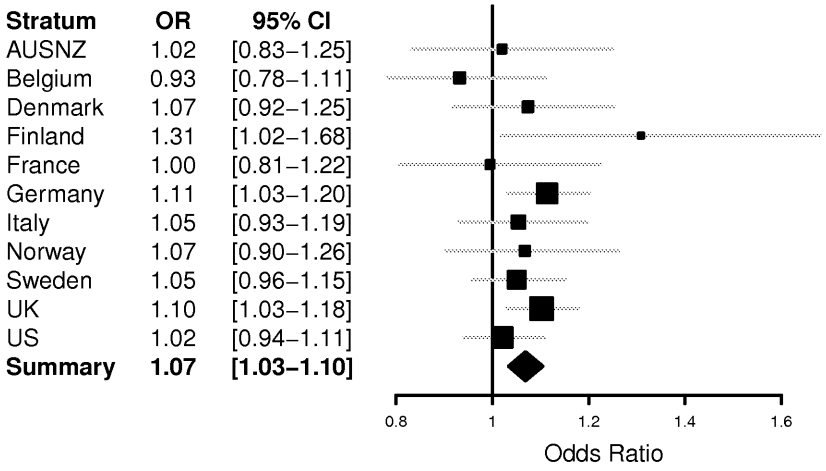


A) Regional Association and B) Forest Plot

Supplementary Figure 94. Discovery phase rs2248359.
A



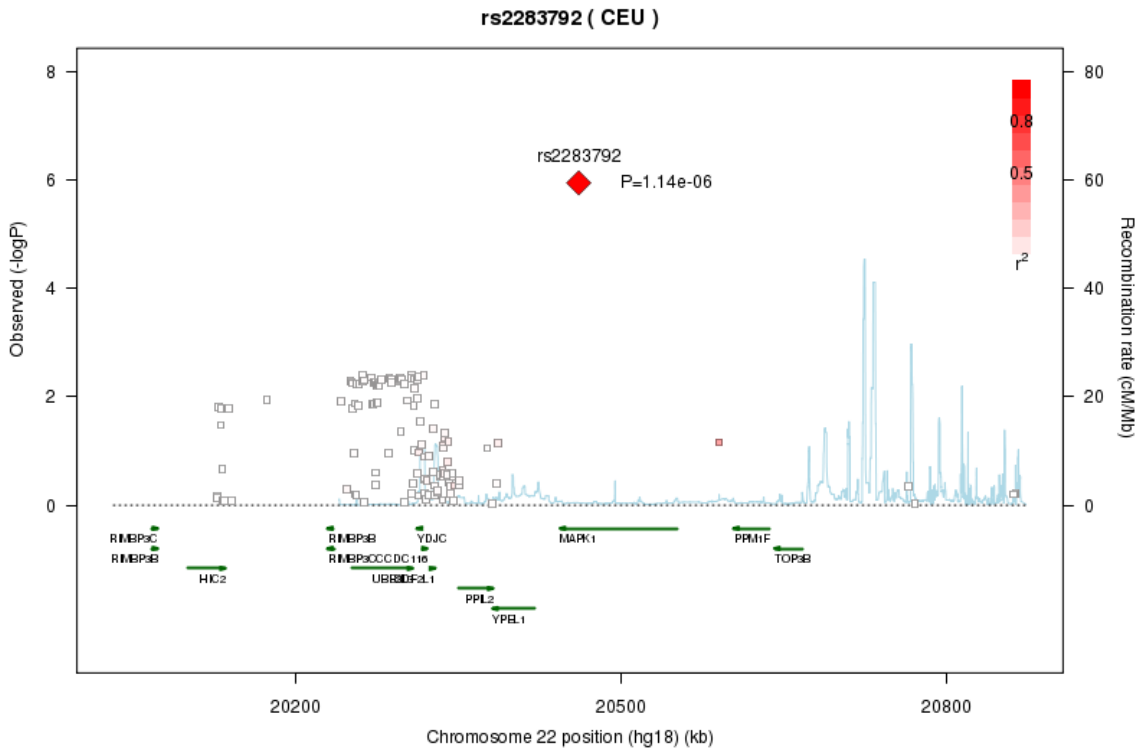
B



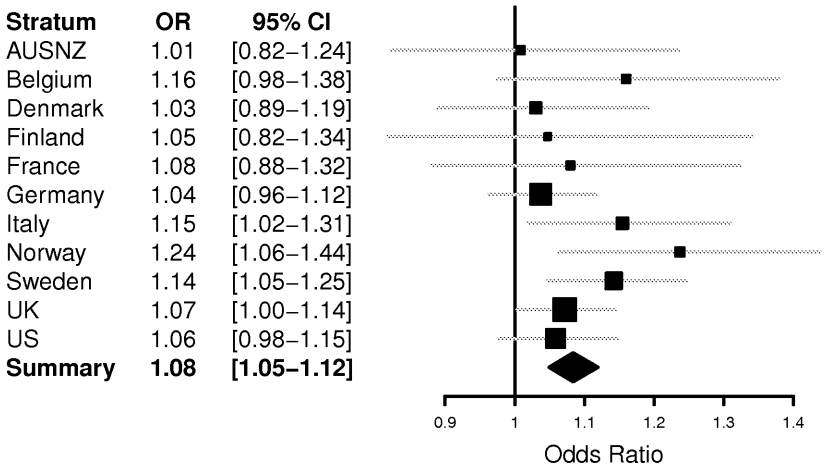
A) Regional Association and b) Forest Plot

Supplementary Figure 95. Discovery phase rs2283792.

A

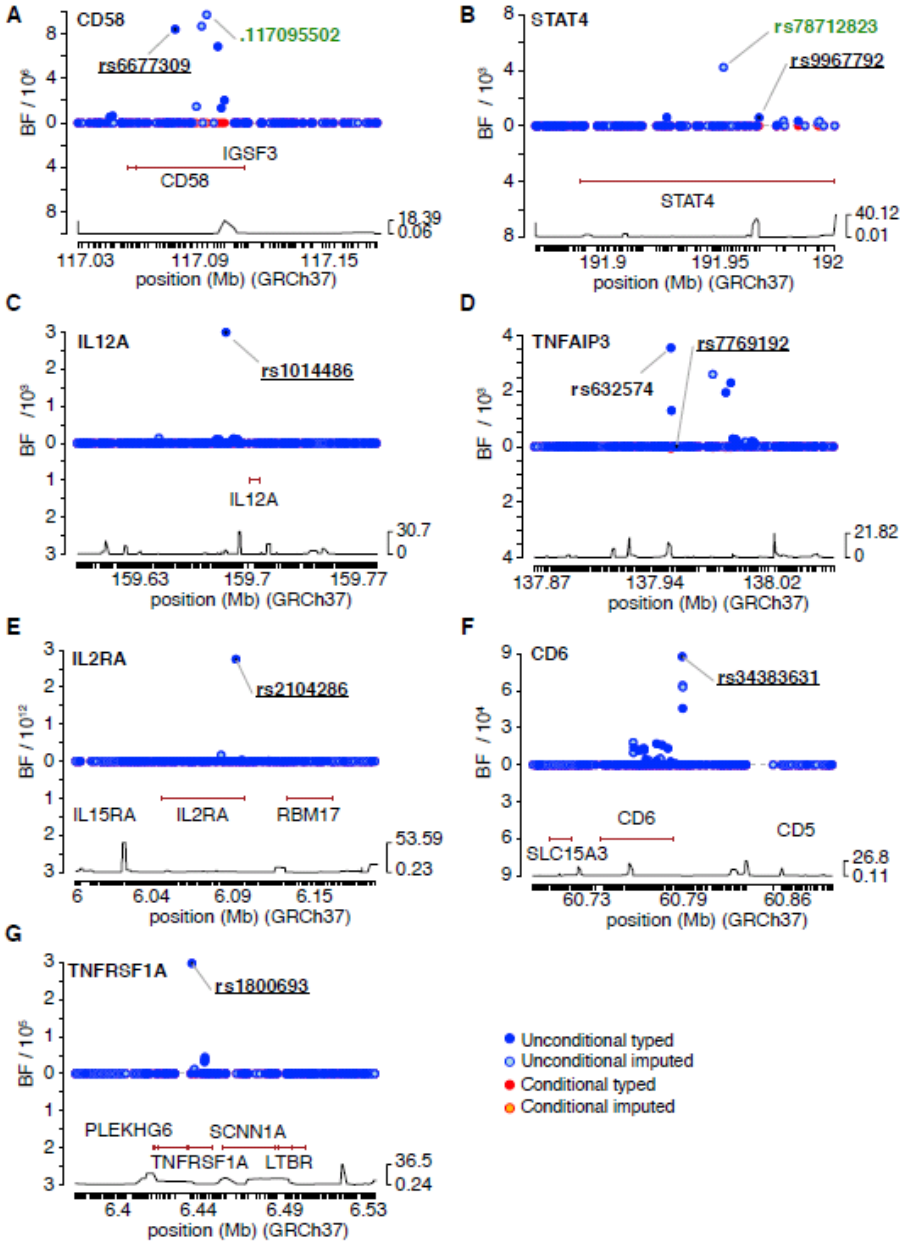


B



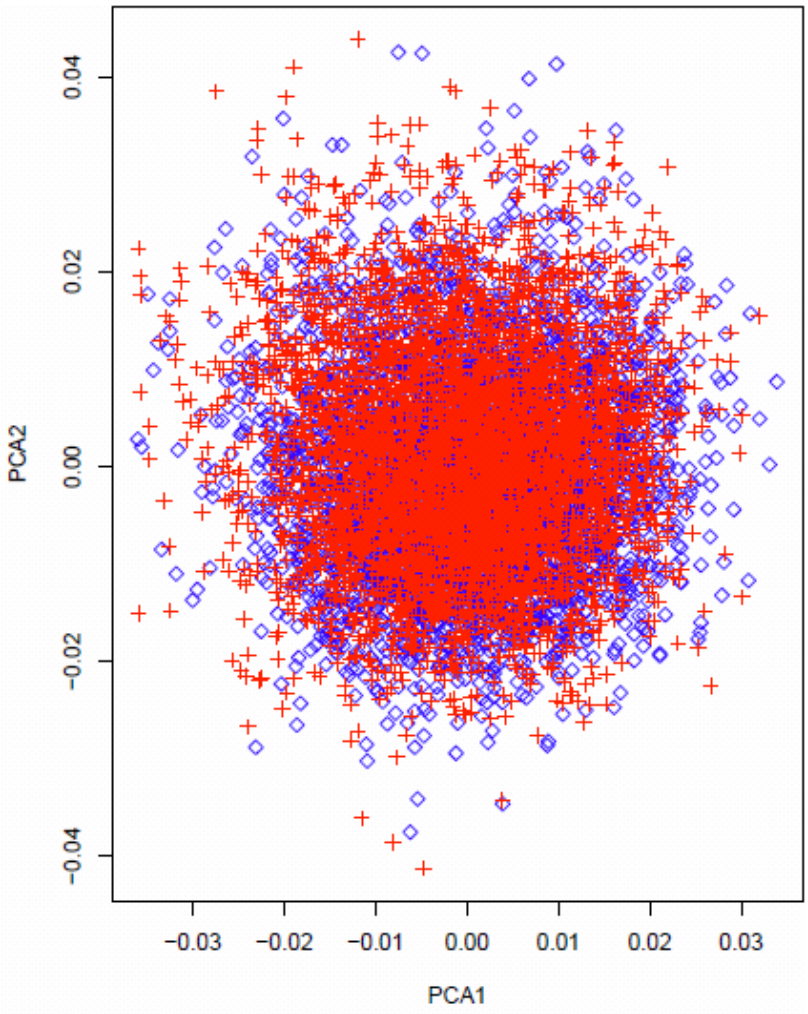
A) Regional Association and B) Forest Plot

Supplementary Figure 96. Regions with Consistent High Resolution Fine-Mapping from Table 3



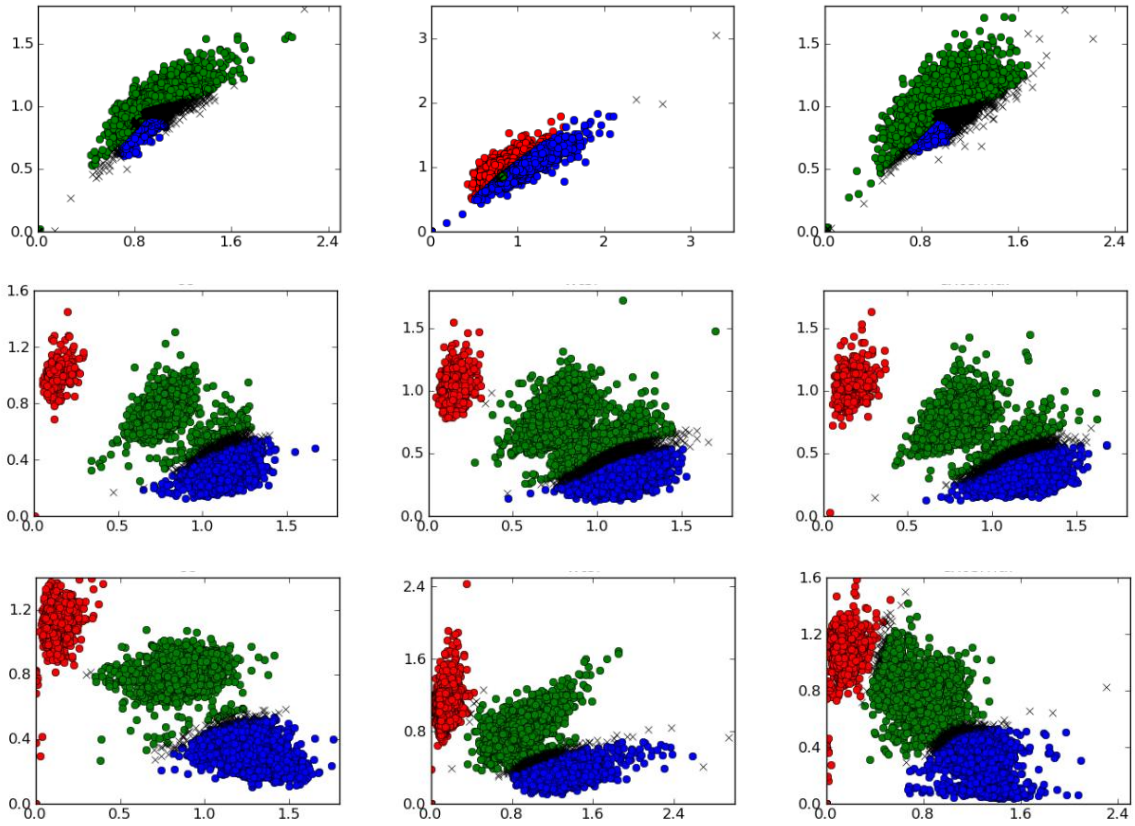
Detail of fine-mapping in regions with consistent high resolution fine-mapping of a) *CD58* b)*STAT4* c)*IL12A* d)*TNFAIP3* e) *IL2RA* f)*CD6* and g) *TNFRSF1A*. Above the x-axis indicates the Bayes Factor summarizing evidence for association for the SNPs prior to conditioning (blue markers) while below the x-axis indicates the Bayes Factor after conditioning on the lead SNP (underlined). In the cases where the variant with maximal Bayes Factor in the UK population was different from the lead SNP identified in the joint discovery plus replication phase, this variant is labelled as well, in green if imputed. Mb=Megabases.

Supplementary Figure 97. The first two PCs of the UK discovery samples.



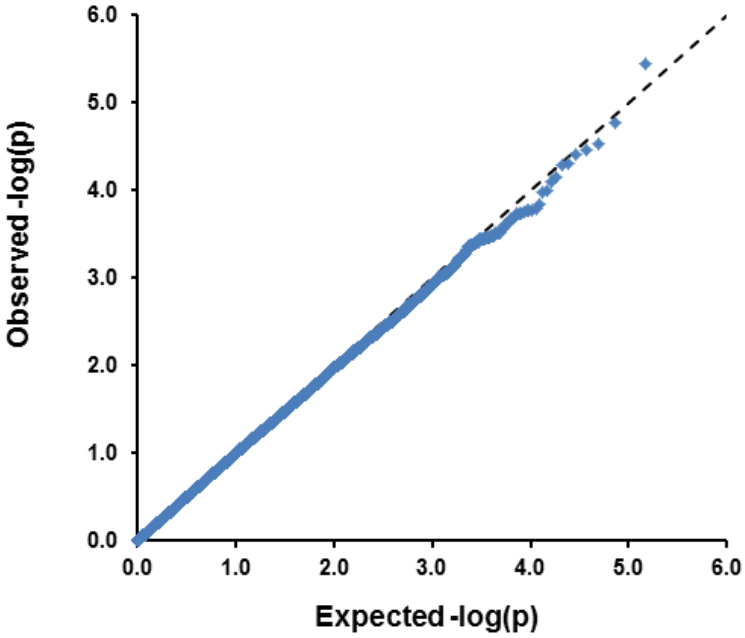
Blue rhomboids depict cases and red crosses depict controls.

Supplementary Figure 98. Examples of cluster plots from SNPs that were excluded on visual inspection.

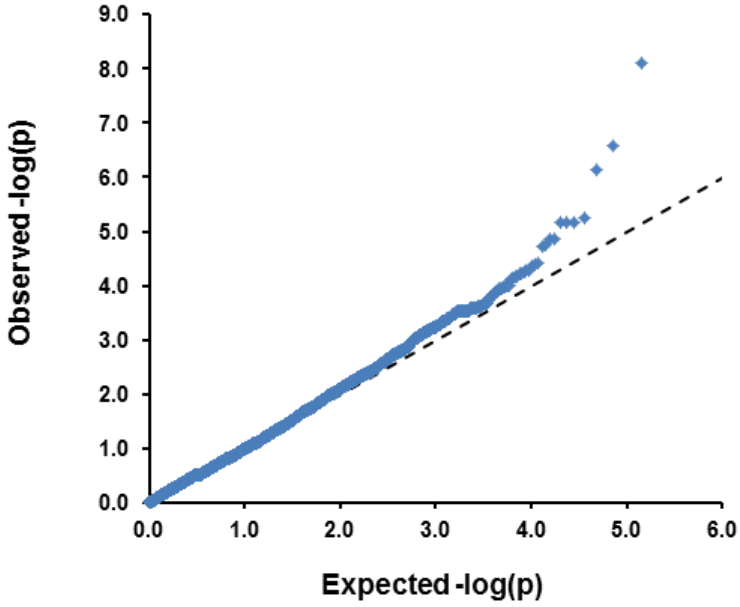


Top row rs9893808, middle row rs3785907 and bottom row rs7206912. For each SNP the left hand panel is based on the WTCCC common controls (n=9999), the centre panel on the multiple sclerosis cases and controls typed at WTSI (n=22837) and the right hand panel on the multiple sclerosis cases and controls that were typed at other centres (n=13049). All plots were generated using Evoker.[33]

Supplementary Figure 99. QQ plot of the MSSS based ImmunoChip analysis of 147136 non-MHC SNPs.



Supplementary Figure 100. QQ plot of the ImmunoChip TDT results for 633 trio families.



References

1. Cortes, A. & Brown, M.A. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* **13**, 101 (2011).
2. IMSGC Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. *N Engl J Med* **357**, 851-62 (2007).
3. Baranzini, S.E. et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* **18**, 767-78 (2009).
4. Bahlo, M. et al. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* **41**, 824-8 (2009).
5. De Jager, P.L. et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* **41**, 776-82 (2009).
6. Refining genetic associations in multiple sclerosis. *Lancet Neurol* **7**, 567-9 (2008).
7. Ban, M. et al. Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. *Eur J Hum Genet* **17**, 1309-13 (2009).
8. IMSGC. The expanding genetic overlap between multiple sclerosis and type I diabetes. *Genes Immun* **10**, 11-4 (2009).
9. IMSGC. Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. *Hum Mol Genet* **19**, 953-62 (2010).
10. IMSGC & WTCCC2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214-9 (2011).
11. IMSGC. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* **70**, 897-912 (2011).
12. IMSGC. MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain* **136,** 1778-82 (2013).
13. Sanna, S. et al. Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nat Genet* **42**, 495-7 (2010).
14. Corrado, L. et al. Association of the CBLB gene with multiple sclerosis: new evidence from a replication study in an Italian population. *J Med Genet* **48**, 210-1 (2011).
15. Poser, C.M. et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* **13**, 227-31 (1983).
16. McDonald, W.I. et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* **50**, 121-7 (2001).
17. Polman, C.H. et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* **58**, 840-6 (2005).
18. Kurtzke, J.F. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* **33**, 1444-52 (1983).
19. Roxburgh, R.H. et al. Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. *Neurology* **64**, 1144-51 (2005).

20. Schumacker, G.A. et al. Problems of Experimental Trials of Therapy in Multiple Sclerosis: Report by the Panel on the Evaluation of Experimental Trials of Therapy in Multiple Sclerosis. *Ann N Y Acad Sci* **122**, 552-68 (1965).
21. Confavreux, C., Compston, D.A., Hommes, O.R., McDonald, W.I. & Thompson, A.J. EDMUS, a European database for multiple sclerosis. *J Neurol Neurosurg Psychiatry* **55**, 671-6 (1992).
22. Lublin, F.D. & Reingold, S.C. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology* **46**, 907-11 (1996).
23. Inouye, M. et al. An immune response network associated with blood lipid levels. *PLoS Genet* **6**(2010).
24. Hedstrom, A.K., Baarnhielm, M., Olsson, T. & Alfredsson, L. Tobacco smoking, but not Swedish snuff use, increases the risk of multiple sclerosis. *Neurology* **73**, 696-701 (2009).
25. Holmen, C. et al. A Swedish national post-marketing surveillance study of natalizumab treatment in multiple sclerosis. *Mult Scler* **17**, 708-19 (2011).
26. Khademi, M. et al. Cerebrospinal fluid CXCL13 in multiple sclerosis: a suggestive prognostic marker for the disease course. *Mult Scler* **17**, 335-43 (2011).
27. Roos, I.M., Kockum, I. & Hillert, J. The interleukin 23 receptor gene in multiple sclerosis: a case-control study. *J Neuroimmunol* **194**, 173-80 (2008).
28. Zajicek, J.P. et al. Patient-orientated longitudinal study of multiple sclerosis in south west England (The South West Impact of Multiple Sclerosis Project, SWIMS) 1: protocol and baseline characteristics of cohort. *BMC Neurol* **10**, 88 (2010).
29. Hirst, C. et al. Increasing prevalence and incidence of multiple sclerosis in South East Wales. *J Neurol Neurosurg Psychiatry* **80**, 386-91 (2009).
30. McWhirter, R.E., McQuillan, R., Visser, E., Counsell, C. & Wilson, J.F. Genome-wide homozygosity and multiple sclerosis in Orkney and Shetland Islanders. *Eur J Hum Genet* **20**, 198-200 (2012).
31. Shah, T.S. et al. optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. *Bioinformatics* **28**, 1598-603 (2012).
32. Teo, Y.Y. et al. A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* **23**, 2741-6 (2007).
33. Morris, J.A., Randall, J.C., Maller, J.B. & Barrett, J.C. Evoker: a visualization tool for genotype intensity data. *Bioinformatics* **26**, 1786-7 (2010).
34. de Bakker, P.I. et al. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum Mol Genet* **17**, R122-8 (2008).
35. Hemminki, K., Li, X., Sundquist, J., Hillert, J. & Sundquist, K. Risk for multiple sclerosis in relatives and spouses of patients diagnosed with autoimmune and related conditions. *Neurogenetics* **10**, 5-11 (2009).
36. Robertson, N.P. et al. Age-adjusted recurrence risks for relatives of patients with multiple sclerosis. *Brain* **119 (Pt 2)**, 449-55 (1996).

37. Willer, C.J., Dyment, D.A., Risch, N.J., Sadovnick, A.D. & Ebers, G.C. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci U S A* **100**, 12877-82 (2003).
38. Zuk, O., Hechter, E., Sunyaev, S.R. & Lander, E.S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A* **109**, 1193-8 (2012).